

Anwendung und Erweiterung der Amber-Stop Codon-Technologie mit Potential zur Generierung biomedizinischer Therapeutika

Carina Andrea Sommer

Vollständiger Abdruck der von der TUM School of Life Sciences der Technischen Universität
München zur Erlangung des akademischen Grades einer

Doktorin der Naturwissenschaften

genehmigten Dissertation.

Vorsitz: Prof. Dr. Mirjana Minceva

Prüfer der Dissertation: 1. Prof. Dr. Arne Skerra
2. Prof. Dr. Wolfgang Liebl

Die Dissertation wurde am 29.08.2023 bei der Technischen Universität München
eingereicht und durch die TUM School of Life Sciences am 28.11.2023 angenommen.

Meiner Schwester Carmen.

Danksagung

Mein besonderer Dank gilt Herrn Prof. Dr. Arne Skerra für die Möglichkeit, ein so spannendes und herausforderndes Thema im Rahmen meiner Doktorarbeit zu bearbeiten. Zudem bedanke ich mich für sein stetes Interesse am Fortgang dieser Arbeit und seinen inspirierenden wissenschaftlichen Ideen.

Ebenso danke ich Dr. Jürgen Lassak für die Unterstützung und für die Übernahme der Mentorenfunktion. Des weiteren möchte ich mich bei Dr. Andreas Eichinger bedanken für Unterstützung bei der kristallographischen Analyse sowie bei Dr. Christopher Graf und Dr. Markus Anneser für die Synthese der nicht-natürlichen Aminosäuren. Außerdem möchte ich mich auch besonders bei Stefan Achatz für die pausenlose Unterstützung bei der massenspektrometrischen Analyse, bei diversen Softwareproblemen und für seinen Humor bedanken.

Hervorheben und danken möchte ich meinem Laborpartner Benjamin Escher, Julia Martin und Jan Scherntaner für die langjährige und freundschaftliche Zusammenarbeit, die gemeinsamen langen Abende und Wochenenden im Labor und der steten Unterstützung sowohl während des Promotionsvorhabens als auch danach.

Bei Kristina Schira bedanke ich mich für die hervorragende Zusammenarbeit und ihren Beitrag zu meinem Promotionsvorhaben im Rahmen ihrer Masterarbeit. Weiters bedanke ich mich bei Firas Hamood für die gemeinsame Arbeit im Rahmen seines Forschungspraktikums.

Herzlicher Dank gilt auch allen meinen langjährigen Mitstreitern im Labor, zum einen für die vielen fachlichen Diskussionen und zum anderen für die unzähligen schönen Momente, die ich mit ihnen erleben durfte. In diesem Rahmen möchte ich mich besonders bei Jonas Schlitz, Sebastian Borgert, Fabian Rodewald, Christian Deuschle, Corinna Brandt, Emanuel Peplau, Laura Job und Elena Ilyukhina bedanken.

Weiterhin großer Dank gilt allen Mitarbeiter des Lehrstuhls für Biologische Chemie, insbesondere Theresia Wiedemann, Martina Dotzauer und Ina Theobald für die unentwegte Unterstützung im Laboralltag und die bestärkenden Gespräche.

Hervorheben und danken möchte ich auch meinen Freunden in Freising und der SGE Eichenfeld-Eittingermoos, ihr hab Freising für mich zum Zuhause gemacht.

Ganz besonderer Dank gilt meiner Familie, meiner ältesten Freundin Chrisi und Paul, ohne euch hätte ich diesen Weg niemals gehen können!

Von ganzem Herzen möchte ich mich bei meiner Freundin Dani für ihre unglaubliche Geduld und Verständnis bedanken. Du hast mich nie daran zweifeln lassen, dass nach dem Regen auch wieder die Sonne scheint.

Teile der vorliegenden Arbeit wurden bereits veröffentlicht:

Edwardraja, S., Eichinger, A., Theobald, I., Sommer, C.A., Reichert, A.J. & Skerra, A. (2017) Rational Design of an Anticalin-Type Sugar-Binding Protein Using a Genetically Encoded Boronate Side Chain. *ACS Synth. Biol* 6, 2241-2247.

Sommer, C.A., Eichinger, A. & Skerra, A. (2020) A Tetrahedral Boronic Acid Diester Formed by an Unnatural Amino Acid in the Ligand Pocket of an Engineered Lipocalin. *ChemBioChem* 21, 469-472.

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Die Erweiterung des genetischen Codes	1
1.2	Neue Anwendungsmöglichkeiten durch die Funktionalisierung von Proteinen mit nicht-natürlichen Aminosäuren	2
1.2.1	Selektive Bindung von Zuckermolekülen durch nicht-natürliche Proteine ...	3
1.2.2	Nicht-natürliche Aminosäuren als spektroskopische Sonden	6
1.3	Die <i>Mj</i> TyrRS: Struktur und Engineering zur selektiven Erkennung von nicht-natürlichen Aminosäuren	7
1.4	Strategien zur Steigerung an Ausbeute an Proteinen mit nicht-natürlichen Aminosäuren	9
1.5	Zielsetzung der Arbeit	12
2	Material und Methoden.....	13
2.1	Material	13
2.1.1	Bakterienstämme und Plasmide.....	13
2.1.2	Gensynthesen und Oligodesoxynukleotide	14
2.1.3	Enzyme und sonstige Proteine.....	17
2.1.4	Chemikalien.....	18
2.1.5	Zucker und Zuckerderivate.....	20
2.1.6	(Nicht-)natürliche Aminosäuren.....	20
2.1.7	Standards und Kits.....	21
2.1.8	Geräte.....	21
2.1.9	Verbrauchsmaterialien	23
2.1.10	Medien, Antibiotika und allgemeine Lösungen	24
2.2	Molekularbiologische Methoden.....	30
2.2.1	Kultivierung und Konservierung von <i>E. coli</i> -Stämmen	30
2.2.2	Transformation von <i>E. coli</i> mit Plasmid-DNA nach der CaCl ₂ -Methode ...	30
2.2.3	Transformation mittels Elektroporation	31
2.2.4	DNA-Isolierung aus <i>E. coli</i>	32
2.2.4.1	Isolierung von chromosomaler DNA.....	32
2.2.4.2	Isolierung von doppelsträngiger Plasmid-DNA	32
2.2.5	Gelelektrophorese und Reinigung von DNA	33
2.2.6	Gen-Deletion mit Hilfe der λ -Red Rekombination	34
2.2.7	Gen-Modifizierung mit Hilfe der Cas9 / λ -Red Rekombination.....	34
2.3	<i>In-vitro</i> -Modifizierung von DNA.....	35
2.3.1	Spaltung doppelsträngiger DNA mit Restriktionsendonukleasen	35
2.3.2	Phosphorylierung und Dephosphorylierung von DNA	36
2.3.3	Polymerase-Kettenreaktion (PCR) zur Amplifizierung von Genen	36
2.3.3.1	PCR-Amplifizierung eines Gens	36
2.3.3.2	Ortsgerichtete Sättigungsmutagenese und <i>Overlap-Extension</i> PCR	37
2.3.3.3	Ortsgerichtete Mutagenese durch QuikChange PCR.....	37
2.3.3.4	Zufallsmutagenese mittels <i>Error Prone</i> PCR.....	38

2.3.3.5	Amplifizierung der <i>Mj</i> TyrRS Zufalls-Bibliothek	38
2.3.3.6	Klonierung einer Gen-Bibliothek mittels Sexual PCR (Gene Shuffling)	39
2.3.4	Ligierung von DNA-Fragmenten	40
2.3.5	Sequenzierung doppelsträngiger DNA	41
2.4	Gentechnische Methoden	41
2.4.1	Produktion rekombinanter Proteine in <i>E. coli</i>	41
2.4.1.1	Periplasmatische Borocalin-Produktion in <i>E. coli</i> im 2 l-Maßstab	41
2.4.1.2	Cytoplasmatische Proteinproduktion in <i>E. coli</i> im 5 ml- und 200 ml-Maßstab	42
2.4.1.3	Vorselektion einer Aminoacyl-tRNA Synthetasen-Bibliothek auf Agarplatten	43
2.5	Proteinchemische Methoden	44
2.5.1	Affinitätschromatographie an Streptavidin-Sepharose	44
2.5.2	Metallchelat-Affinitätschromatographie	44
2.5.3	Ionenaustauschchromatographie	45
2.5.4	Größenausschlusschromatographie	45
2.5.5	SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE)	46
2.5.6	Konzentrierung von Proteinen	46
2.5.7	Bestimmung von Proteinkonzentrationen	47
2.6	Biophysikalische Methoden	47
2.6.1	Fluoreszenzspektroskopie	47
2.6.2	Massenspektrometrie	48
2.6.3	Cytofluorimetrische Selektion aus Gen-Bibliotheken und Einzelklonanalyse mittels <i>Fluorescence Activated Cell Sorting</i> (FACS) ...	49
2.7	Proteinkristallographische Methoden	50
2.8	Verwendete Computerprogramme und Datenbanken	51
3	Ergebnisse	53
3.1	Engineering eines Lipocalins mit einer nicht-natürlichen Aminosäure für die Bindung von Zuckermolekülen	53
3.1.1	Produktion und Reinigung rekombinanter Borocalin-Varianten	55
3.1.2	Evaluierung der Bindungsaffinität von Borocalin-Varianten gegenüber Zuckern und Zuckerderivaten	56
3.1.3	Kristallstruktur der Borocalin-Variante Lcn2a36Bpa-NFW im Komplex mit 4-Nitrocatechol	61
3.1.4	Rationales Engineering ausgehend von Lcn2a36Bpa-NFW	64
3.1.5	Eine reaktivere Bor-Aminosäure für die selektive Bindung von Zuckermolekülen	68
3.2	Engineering und Selektion einer auf der Tyrosyl-tRNA Synthetase aus <i>Methanocaldococcus jannaschii</i> basierenden Aminoacyl-tRNA Synthetase für den Einbau von 7-Hydroxycoumarinalanin (Hca)	70
3.2.1	Die SlonoMax Genbibliothek	71
3.2.1.1	Selektion aus der SlonoMax Genbibliothek	71

3.2.1.2	Selektion aus der SlonoMax Genbibliothek nach einer Vorselektion	74
3.2.2	Selektion einer Aminoacyl-tRNA Synthetase einer auf Substratgruppen fokussierten Zufalls-Bibliothek der <i>MjTyrRS</i>	81
3.2.2.1	Strukturelle Analyse evolvierter Aminoacyl-tRNA Synthetasen ...	81
	Structure of AzuAla.....	83
3.2.2.2	Konstruktion einer auf Substratgruppen fokussierten Zufalls-Bibliothek mittels Sexual PCR.....	85
3.2.2.3	Einzelklonanalyse und Sequenzvergleich der selektierten Aminoacyl-tRNA Synthetasen zur molekularen Erkennung von Hca.....	89
3.3	Unterdrückung der Konkurrenz des Release Faktors 1 um das Amber Stopp Codon zur Steigerung der Ausbeute an rekombinantem eGFP(39a) mit einer nicht-natürlichen Aminosäure	91
3.3.1	Deletion des Release Faktors 1 in einem <i>E. coli</i> B-Stamm.....	91
3.3.1.1	Cytofluorimetrische Analyse und Vergleich der Expression von eGFP(39a) in den Stämmen <i>E. coli</i> NEB Express Δ Rf-1 und <i>E. coli</i> NEB Express	94
3.3.1.2	Produktion von eGFP(39Bpa) in den Stämmen <i>E. coli</i> NEB Express Δ Rf-1 und <i>E. coli</i> NEB Express, Reinigung und quantitativer Vergleich der Ausbeuten.....	95
3.3.2	Fakultative Inaktivierung des Release Faktors 1 in einem <i>E. coli</i> K12-Stamm.....	96
3.3.2.1	Einführung der Erkennungssequenz der TEV-Protease mittels Cas9 und λ -Red Rekombination in den Switch Loop des Release Faktors 1.....	96
3.3.2.2	Cytofluorimetrische Analyse und Vergleich der <i>E. coli</i> MG1655-Tev Varianten hinsichtlich der Expression von eGFP(39a)	100
3.3.2.3	Cytofluorimetrische Analyse und Vergleich der Expression von eGFP(39a) mit den Stämmen <i>E. coli</i> MG1655-Tev2a und <i>E. coli</i> MG1655	103
3.3.2.4	Expression von eGFP(39a) mit den Stämmen <i>E. coli</i> MG1655-Tev2a und <i>E. coli</i> MG1655, Reinigung und quantitativer Vergleich der Ausbeuten	107
3.3.2.5	Fusion von TEV-Protease mit dem Sumo-Protein zu Steigerung der Löslichkeit Aktivität.....	108
3.3.2.6	Evaluierung der maximal erreichbaren Suppression des Amber-Stop Codons mittels der <i>E. coli</i> Suppressor-tRNA SupE	110
4	Diskussion.....	114
4.1	Synthetische Biologie: Die Zukunft des Protein-Engineerings.....	114
4.2	Borocaline als Biologika für die Diagnostik und Therapie von Krebserkrankungen und Virusinfektionen	115
4.3	Herausforderung bei der Selektion spezifischer Aminoacyl-tRNA Synthetasen für nicht-natürliche Aminosäuren	117

4.4	Limitierende Faktoren für die Proteinausbeute auf dem Weg zur Anwendbarkeit	119
5	Zusammenfassung.....	123
6	Abkürzungsverzeichnis	125
7	Literaturverzeichnis	127

1 Einleitung

1.1 Die Erweiterung des genetischen Codes

Der genetische Code gilt als zentrales Dogma für alle Formen des Lebens und verbindet die genetische Information der DNA beziehungsweise nach der Transkription die genetische Information der mRNA mit der Proteinbiosynthese (Crick, 1958; Cobb, 2017). Er basiert auf der Gruppierung von drei aufeinanderfolgenden Nukleinsäuren zu sogenannten Codons, welche jeweils einer bestimmten Aminosäure zugeordnet sind (Woese, 1964; Hinegardner & Engelberg, 1963). Aus den vier Ribonukleotiden Adenosin, Guanosin, Cytidin und Thymin beziehungsweise Uridin in der mRNA ergeben sich 64 mögliche Codons. Drei dieser Codons sind Stop Codons, an denen die Translation terminiert wird und die historisch bedingt Amber-(TAG), Ocker- (TAA), oder Opal-Stop Codon (TGA) genannt werden. Für die 20 proteinogenen Aminosäuren kodieren die übrigen 61 Codons, deshalb wird der genetische Code auch als degeneriert bezeichnet. Dabei ist die Verteilung der Aminosäuren auf die zur Verfügung stehenden Codons nicht zufällig (Woese, 1965; Crick, 1968).

Bei näherer Betrachtung fällt auf, dass Codons, die für dieselbe Aminosäure oder für Aminosäuren, die ähnliche chemische Eigenschaften aufweisen, kodieren, sich nur in einer Base unterscheiden (Abbildung 1).

		Zweite Base									
		U		C		A		G			
Erste Base	U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U	Dritte Base
		UUC		UCC		UAC		UGC		C	
		UUA	Leu	UCA		UAA	End	UGA	End	A	
		UUG		UCG		UAG	End	UGG	Trp	G	
	C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U	
		CUC		CCC		CAC		CGC		C	
		CUA		CCA		CAA	CGA	A			
		CUG		CCG		CAG	CGG	G			
	A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U	
		AUC		ACC		AAC		AGC		C	
		AUA		ACA		AAA	AGA	A			
		AUG	Met	ACG		AAG	Lys	AGG	Arg	G	
	G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U	
		GUC		GCC		GAC		GGC		C	
		GUA		GCA		GAA	GGA	A			
		GUG		GCG		GAG	GGG	G			

Abbildung 1: Der genetische Code.

Im Prinzip könnte der genetische Code zusätzliche Informationen für weitere Aminosäuren speichern. Diese Möglichkeit wird in der Natur jedoch kaum genutzt (Ambrogelly *et al.*, 2007). Die natürliche Erweiterung des genetischen Codes beschränkt sich auf die beiden Aminosäuren Selenocystein und Pyrrolysin, die deshalb auch als die 21. und 22. Aminosäure bezeichnet

werden. Der Ko-translationelle Einbau von Selenocystein in das Protein Formiat-Dehydrogenase wurde 1988 entdeckt (Zinoni *et al.*, 1987; Leinfelder *et al.*, 1988). Dabei wurde ersichtlich, dass für den Einbau dieser lebensnotwendigen Aminosäure, der kein eigenes Codon zugeordnet ist, das Opal-Stop Codon rekodiert und dadurch die an diesem Codon vorgesehene Termination supprimiert wird. Diese Vorgehensweise findet man in beinahe allen Formen des Lebens (Forchhammer & Böck, 1991). Zehn Jahre nach diesem Fund wurde ein einzelnes Amber-Stop Codon in dem Leseramen des Gens für eine Monomethylamin Methyltransferase aus *Methanosarcina bakeri* entdeckt, dass die Translation jedoch nicht terminiert, (Burke *et al.*, 1998), stattdessen wurde an dieser Position Pyrrolysin in die Polypeptidkette eingebaut (Hao *et al.*, 2002).

Diese Vorgehensweise der Natur kann als Blaupause genutzt werden, um gezielt modifizierte Aminosäuren-Seitenketten in gentechnisch hergestellte Proteine einzubauen und dadurch das Set der 20 natürlichen Aminosäuren um zusätzliche physikalisch-chemische Eigenschaften zu erweitern (Wals & Ovaas, 2014; Young & Schultz, 2010). Bestrebungen dahingehend existierten schon seit den frühen 2000er Jahren (Wang *et al.*, 2000) und umfassen folgende Maßnahmen: Erstens muss ein Codon gewählt werden, das sich zur Rekodierung eignet. Dabei bietet sich vor allem das Amber-Stop Codon an, das am seltensten genutzte Stop Codon, oder andere seltene, für Aminosäuren kodierende Codons (Zeng *et al.*, 2014). Zweitens benötigt man ein orthogonales Paar aus Aminoacyl-tRNA Synthetase und Suppressor-tRNA. Orthogonal bedeutet, dass die Aminoacyl-tRNA Synthetase keine endogene tRNA erkennt bzw. dass die Suppressor-tRNA der gewählten Wirtszelle auch nicht von einer endogenen Aminoacyl-tRNA Synthetase als Substrat genutzt wird. Deshalb verwendet man abhängig von der gewählten Wirtszelle ein Paar aus Aminoacyl-tRNA Synthetase und Suppressor-tRNA aus einem nicht verwandten Organismus. Für die gentechnische Produktion von rekombinanten Proteinen mit *E. coli* kommen Aminoacyl-tRNA Synthetasen aus Archaeobakterien zum Einsatz, zum Beispiel eine engierte TyrRS aus *Methanocaldococcus jannaschii* (Wang *et al.*, 2001) oder eine engierte PylRS aus *Methanosarcina barkeri* (Wan *et al.*, 2014). Diese engierte Aminoacyl-tRNA Synthetase erkennt das Aminosäure-Substrat spezifisch, adenyliert und lädt dieses auf die Suppressor-tRNA (Pang *et al.*, 2014). Während der ribosomalen Proteinbiosynthese erkennt die Suppressor-tRNA mit ihrem Anticodon das Amber-Stop Codon auf der mRNA und die Aminosäure wird in die Polypeptidkette eingebaut.

1.2 Neue Anwendungsmöglichkeiten durch die Funktionalisierung von Proteinen mit nicht-natürlichen Aminosäuren

Das Forschungsfeld um den Einbau nicht-natürlicher Aminosäuren in Proteine trägt zum einen in der Grundlagenforschung zur Aufklärung von biologischen Prozessen bei. So ermöglichte es diese Technologie durch den Einbau von *p*-Benzoyl-L-phenylalanin und die dadurch erzeugten

Protein-Quervernetzungen das Verhalten von Proteinen *in vivo* mit zu analysieren (Davis & Chin, 2012). Des Weiteren ist es mit fluoreszierenden nicht-natürlichen Aminosäuren möglich sowohl Proteinkonformationen durch Förster-Resonanzenergietransfer (FRET) (Wang *et al.*, 2014) als auch die intrazelluläre Proteinlokalisierung aufzuklären (Neumann, 2012). Der Einbau von nicht-natürlichen Aminosäuren in Proteine, die ein Schwermetall beziehungsweise ein Fluoratom beinhalten, erleichtern die Bestimmung der Proteinstruktur mittels Röntgenkristallographie oder ^{19}F NMR Spektroskopie (Seifert *et al.*, 2002; Ai, 2012). Zum anderen erweitert die Funktionalisierung von Proteinen mit nicht-natürlichen Aminosäuren das Anwendungsspektrum durch die daraus resultierenden und mit neuen chemischen Eigenschaften ausgestatteten synthetischen Proteinen. Dabei kann der Einbau einer nicht-natürlichen Aminosäure die physikalischen und chemischen Eigenschaften von Proteinen wie zum Beispiel den isoelektrischen Punkt, Polarität, Stabilität sowie spektroskopische Eigenschaften verändern (Budisa, 2006). Die wohl bekannteste Funktionalisierung von Proteinen dieser Art erfolgte durch das Einbringen von nicht-natürlichen Aminosäuren für eine bioorthogonale Click-Reaktion (Prescher & Bertozzi, 2005). Das ermöglichte sowohl die Proteinimmobilisierung als auch die positionsspezifische Derivatisierung der Proteine durch funktionelle Gruppen wie zum Beispiel Polyethylenglycol, Zucker, Oligonukleotide oder Fluorophore (Lang *et al.*, 2015). Nicht-natürliche Aminosäuren mit photoaktiven Gruppen erlauben die gesteuerte Induktion der durch die nicht-natürliche Aminosäure vermittelten Eigenschaft aufgrund gezielter Freisetzung einer Schutzgruppe (Neumann, 2012) oder einer durch UV-Licht vermittelten Konformationsänderung der nicht-natürlichen Aminosäure (Skerra *et al.*, 2018).

1.2.1 Selektive Bindung von Zuckermolekülen durch nicht-natürliche Proteine

Eine weitere Anwendungsmöglichkeit von nicht-natürlichen Aminosäuren ist die Funktionalisierung von Bindeproteinen mit einer borhaltigen nicht-natürlichen Aminosäure (*p*-Borono-L-phenylalanin, Bpa) für die chemische Komplexbildung mit Kohlenhydraten. (Edwardraja *et al.*, 2017). Kohlenhydrat/Protein-Interaktionen spielen eine wichtige Rolle in vielen biologischen oder pathologischen Prozessen, wie zum Beispiel bei Zell/Zell-Adhäsion, Entzündungen, Metastasierung von Krebszellen, Befruchtung der Eizelle, Diabetes und Virusinfektionen (Varki, 2016; Reily *et al.*, 2019). Dabei bilden beteiligte Kohlenhydratstrukturen ein attraktives Ziel sowohl für die Diagnostik als auch für therapeutische Interventionen. Bei Virusinfektionen kann durch Blockieren der Zuckerbindung die Fusion des Virus mit der Wirtszelle unterbunden werden und Glycanstrukturen wie die Lewis-Antigene werden als Biomarker für Krebs verwendet (Blanas *et al.*, 2018).

Die Affinität der natürlichen Kohlenhydratbindepoteine, sogenannter Lektine, erweist sich generell als schwach, sodass in der Regel eine multivalente Interaktion erforderlich ist (Ambrosi *et al.*, 2005). Die monovalente Affinität liegt bei Monosacchariden in einem K_D -Wert-Bereich von mM bis nM (Toone, 1994). Die Bindung von Lektinen an Kohlenhydrate erfolgt sowohl über Wasserstoffbrücken zwischen den polaren Aminosäureseitenketten des Proteins und den Hydroxylgruppen der Zuckerderivate als auch vermittelt durch Metallionen. Des weiteren interagieren Lektine über hydrophobe Wechselwirkungen zwischen unpolaren Aminosäureseitenketten und Methin-Gruppen (CH-Gruppen) des Zuckers, wobei aromatische Aminosäuren durch CH- π Wechselwirkungen involviert sein können (Gabijs *et al.*, 2011; Weis & Drickamer, 1996). Die geringe Affinität der Kohlenhydrat-Bindeproteine liegt zum einen daran, dass die Struktur der Kohlenhydrate von Hydroxylgruppen dominiert wird und diese den Wassermolekülen des Solvent ähnlich sind. Zum anderen unterscheidet sich die Struktur von Oligosacchariden aufgrund der freien Bindungswinkel an der glycosidischen Bindung stark voneinander (Laine, 1994), das in einem Verlust der inneren Entropie bei der Komplexbildung mit einem Protein resultiert.

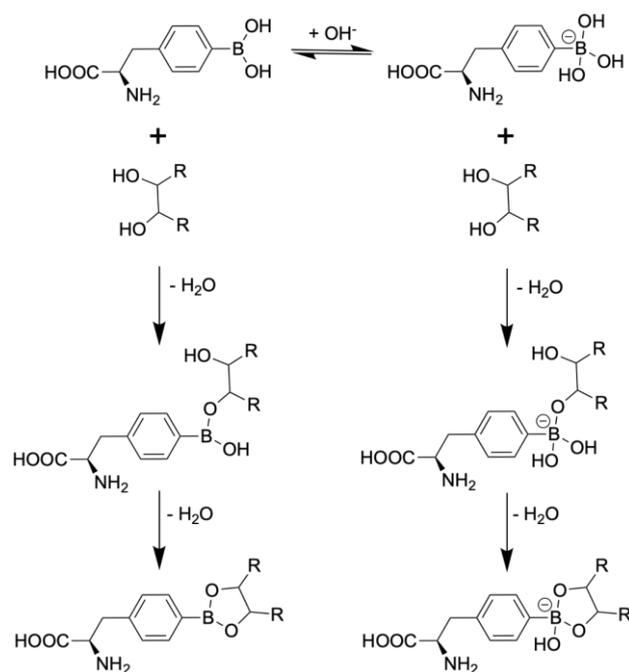


Abbildung 2: pH-abhängiges chemisches Gleichgewicht der Reaktion nicht-natürlichen Aminosäuren Bpa mit einem *cis*-Diol. Mono-Esterbindung und weitere Kondensierung zu einem entropisch bevorzugten zyklischen di-Ester, das in der trigonal-planaren Struktur (links) oder in der tetraedrischen Struktur (rechts) der Boronsäure resultiert.

Die Generierung von hochaffinen Bindeproteinen zur selektiven Bindung von Glycanstrukturen oder Monosacchariden unter physiologischen Bedingungen für medizinische Diagnostik und Behandlung ist daher ein aktuelles Forschungsgebiet. Im Mittelpunkt steht derzeit die Generierung von synthetischen Lektinen, die den natürlichen Zuckerbindungspartnern nachempfunden sind (Davis, 2009; Kubik, 2009; Duan *et al.*, 2020) oder Antikörper gegen

Zuckerstrukturen (Kappler & Hennet, 2020a). Das grundlegende Problem, die Ähnlichkeit der Zuckermoleküle zum wässrigen Lösungsmittel, bleibt jedoch persistent. Der Einsatz von Boronsäuren bietet hier im Prinzip einen Ausweg. Borsäure ist seit langem dafür bekannt, zyklische Diester mit *cis*-Diolen zu bilden, wie sie auch in vielen Zuckermolekülen zu finden sind (Peters, 2014; v. Vargha, 1933). Im trigonal-planaren Grundzustand weisen Boronsäuren ein leeres Valenzorbital auf und verhalten sich als Lewis-Säuren, die beispielsweise im Fall von Phenylboronsäuren, bei einem hohen pH-Wert mit einer Lewis-Base (OH⁻) einen anionischen Tetraeder bilden (Pizer & Babcock, 1977). Dem gegenüber läuft die Reaktion von Boronsäuren mit organischen 1,2 *cis*-Diolen in einem zweistufigen Prozess ab, der pH-abhängig in einem trigonal-planaren oder in einem tetraedrischen Komplex resultiert (Abbildung 2) (Peters, 2014).

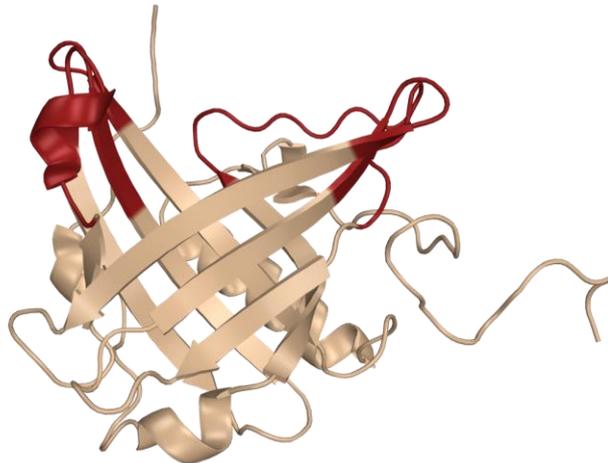


Abbildung 3: Struktur des humanen Lipocalin 2 (Lcn2). Die kelchartige Form wird von 8 antiparallelen β -Faltblattsträngen in β -Fass-Anordnung und einer anliegenden α -Helix geformt. Die hier in Rot dargestellten variablen Schleifen (1-4) bilden die natürliche Ligandenbindestelle aus.

Der Einbau einer borhaltigen nicht-natürliche Aminosäuren, Bpa, in das humane Lipocalin 2 (Lcn2, UniProt ID: P80188) ermöglicht die Funktionalisierung dieses Bindeproteins zur reversiblen Komplexbildung mit *cis*-Diolen (Edwardraja *et al.*, 2017). Die Lipocaline sind eine Klasse von kleinen Bindeproteinen (150 – 180 Aminosäuren), welche in nahezu allen Lebewesen zu finden sind (Åkerström *et al.*, 2006). Allein im menschlichen Körper befinden sich mehr als 12 Vertreter der Lipocalin-Familie, deren physiologische Funktion in der Speicherung, dem Transport oder der Ausscheidung von schädlichen Substanzen liegt, wie zum Beispiel Pheromone, Steroide und Vitamine. Die Lipocalin-Familie zeichnet sich durch eine geringe Homologie der Primärsequenz aber eine ausgeprägte Ähnlichkeit in der Tertiärstruktur aus, die auf einer Kelchstruktur basiert (Schiefner & Skerra, 2015). Geformt wird dieser Kelch von 8 antiparallelen β -Faltblattsträngen in β -Fass-Anordnung (β -Barrel) und einer anliegenden α -Helix. Eine Seite des β -Fasses ist durch enge Peptidschleifen mit hydrophoben Aminosäure-Seitenketten verschlossen. Die gegenüberliegende, offene Seite des β -Fasses weist jedoch vier

strukturell variable Schleifen auf, die die β -Faltblattstränge ebenfalls paarweise miteinander verbinden und dabei die natürliche Ligandenbindestelle ausbilden (Abbildung 3) (Skerra, 2000). Diese strukturelle Eigenschaft wird bei der Anticalin-Technologie genutzt, um durch kombinatorisches Protein-Design künstliche Ligandenbindestellen vor allem ausgehend von Lcn1 oder Lcn2 zu generieren (Gebauer *et al.*, 2013; Richter *et al.*, 2014). Auf diese Weise konnte bereits eine Reihe vielversprechender Anticaline für biomedizinische Anwendungen entwickelt werden (Rothe & Skerra, 2018). Daher bietet das Lcn2-Grundgerüst auch eine ideale Grundlage für die Funktionalisierung mit einer Boronsäure, um gegebenenfalls anschließend durch weiteres Modifizieren der Bindetasche selektive Zuckerbindung zu ermöglichen und einen hochaffinen, synthetischen Kohlenhydrat-Rezeptor zu generieren.

1.2.2 Nicht-natürliche Aminosäuren als spektroskopische Sonden

Die Fluoreszenzmarkierung von Proteinen ist in der Grundlagenforschung und in der biomedizinischen Diagnostik bei der Aufklärung von Protein-Interaktionen von Bedeutung. Markiert werden die Proteine meist *in vitro*, an Aminosäure-Seitenketten der Proteinoberfläche mit einem Fluoreszenzfarbstoff oder *in vivo* durch Fusion am C- oder N-Terminus mit einem fluoreszierenden Protein oder Peptid (Toseland, 2013). Diese Arten der Fluoreszenzmarkierung gehen aber mit erheblichen strukturellen Modifikationen einher und können dadurch die biomolekularen Eigenschaften der markierten Peptide oder Proteine beeinflussen. Um einen Fluoreszenzfarbstoff an einer bestimmten Position, zum Beispiel in der Nähe der Bindetasche, zu platzieren, können alternativ fluoreszierende nicht-natürliche Aminosäuren mittels Amber-Stop Codon Suppression in die Polypeptidkette eingebaut werden (Cheng *et al.*, 2020). Das wurde unter anderem bereits mit den nicht-natürlichen Aminosäuren Dansylalanin (Wang *et al.*, 2006), Hydroxycoumarylethylglycin (Summerer *et al.*, 2006), Hydroxyquinolinalanin (Liu *et al.*, 2013) und Acetylnaphthalylalanin (Chatterjee *et al.*, 2013) publiziert. Allerdings sind diese nicht-natürlichen Aminosäuren sterisch anspruchsvoll und würden bei einem gezielten Einbau in eng gepackte hydrophobe Bereiche eines Proteins die Proteinpäckung beeinflussen. Daher eignen sie sich kaum zur Funktionalisierung eines Bindeproteins tief in der Bindetasche, um Interaktionen spektroskopisch nachzuweisen.

Eine nicht-natürliche Aminosäure, die sich für derlei Anwendungen besser eignen würde, ist 7-Hydroxycoumarylalanin (Abbildung 29), welche im Gegensatz zu Hydroxycoumarylethylglycin (Wang *et al.*, 2006; Kuhn *et al.*, 2011) eine weniger lange und flexible Aminosäure-Seitenkette aufweist. Coumarin und seine Derivate sind bemerkenswerte fluoreszierende Moleküle. Trotz der geringen Molekülgröße weisen sie eine relativ hohe Quantenausbeute auf sowie einen geringen *Bleaching*-Effekt, zeigen hohe Stabilität und sind leicht zu synthetisieren (Brun *et al.*, 2004). Deshalb wurden Coumarin-Derivate — neben der verbreiteten Anwendung als fluorogene Abgangsgruppe in Enzymsubstraten — schon häufig zur Fluoreszenzmarkierung für experimentelle Anwendungen genutzt (Carneiro *et al.*, 2021).

1.3 Die *Mj*TyrRS: Struktur und Engineering zur selektiven Erkennung von nicht-natürlichen Aminosäuren

Wie bereits in Abschnitt 1.1 erwähnt, benötigt man für den spezifischen Einbau einer nicht-natürlichen Aminosäure durch Amber-Stop Codon Suppression ein jeweils für die gewählte nicht-natürliche Aminosäure spezifisches Paar aus Aminoacyl-tRNA Synthetase und tRNA aus einem nicht verwandten Organismus gewählt, um Orthogonalität zu gewährleisten (Melnikov & Söll, 2019). Bewährt für den Einbau von nicht-natürlichen Aminosäuren in rekombinante Proteine in *Escherichia coli* (*E. coli*) haben sich: TyrRS/tRNA aus *Methanocaldococcus jannaschii* für sterisch anspruchsvolle und PylRS/tRNA aus *Methanosarcina barkeri* oder *Methanosarcina mazei* für langgestreckte Aminosäure-Seitenketten.

Da das Paar *Mj*TyrRS/tRNA für diese Arbeit von größerer Bedeutung ist, wird deren Struktur und Funktion hier genauer beschrieben. Die für Tyrosin spezifische *Mj*TyrRS ist eine Aminoacyl-tRNA Synthetase der Klasse I und liegt als Homo-Dimer vor, wobei jedes Monomer aus 5 Domänen besteht. Die Struktur beginnt am N-Terminus mit einer kurzen α -helicalen Domäne gefolgt von einer Rossman-Fold-Domäne, die die Aminosäure erkennt und ein kurzes Verbindungspeptid beinhaltet (CP1), das für die Dimerisierung das *Interface* beider Homo-Dimere bildet. Den Abschluss bildet eine als KMSKS-Schleife bekannte Region, beziehungsweise in dieser Klasse eine KMSSS-Schleife, die die Rossman-Fold-Domäne mit dem C-terminalen Domäne verbindet (Kobayashi *et al.*, 2003). Der *Mj*TyrRS fehlt eine Korrektur-Domäne, um die Beladung der Suppressor-tRNA nochmals auf ihr Korrektheit zu prüfen, und eine Anticodon-Bindedomäne (Jakubowski & Goldman, 1992). Ohne diese beiden Domänen lässt sich sowohl die Substratspezifität als auch das Anticodon der dazugehörigen tRNA verändern.

Hinsichtlich der Orthogonalität des *Mj*TyrRS/tRNA Paares ist entscheidend, wie diese Suppressor-tRNA von der *Mj*TyrRS erkannt wird. Dabei ist das erste Basenpaar (C1:G72) des Akzeptorstammes am bedeutendsten und unterscheidet sich von der bakteriellen Tyrosyl-tRNA (G1:C72). Die Base C1 der Suppressor-tRNA ist um 20 ° gedreht, wodurch die O2 und N3 Atome dieser Base mit Arg¹⁷⁴ Wasserstoffbrücken ausbilden. Des weiteren interagiert C1 über hydrophobe Wechselwirkungen mit Met¹⁷⁸. Die Base C1 trägt zwar hauptsächlich zur molekularen Erkennung der Suppressor-tRNA durch die *Mj*TyrRS bei, aber auch die Base G72 interagiert über Wasserstoffbrücken mit Lys¹⁷⁵ (Kobayashi *et al.*, 2003). Obwohl angenommen wurde, dass die Suppressor-tRNA aufgrund der Basenpaarung C1:G72 in *E. coli* orthogonal ist, wurde gezeigt, dass endogene Aminoacyl-tRNA Synthetase diese tRNA, wenn auch nur schlecht, erkennen können. Daher wurde eine optimierte Suppressor-tRNA generiert und die Erkennung durch endogene Aminoacyl-tRNA Synthetasen um den Faktor 4 reduziert (Wang *et al.*, 2001).

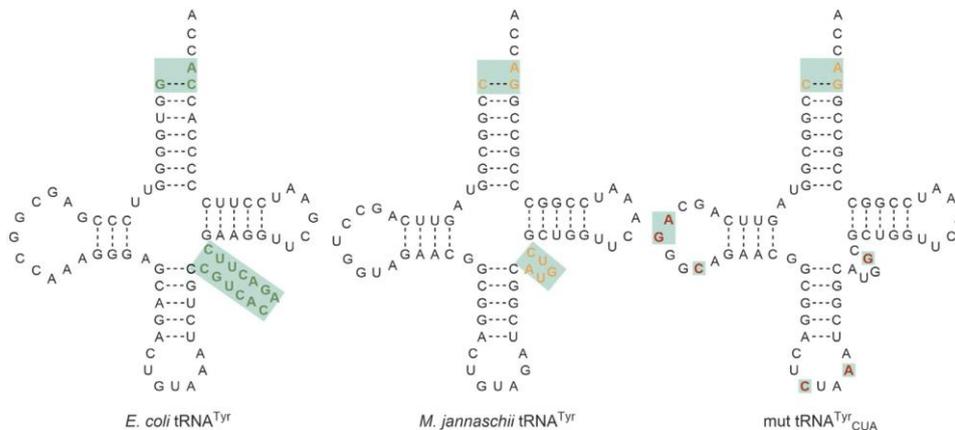


Abbildung 4: Schematische Darstellung von tRNAs, welche für die Aminosäure Tyr kodieren. Endogene Tyrosyl-tRNA von *E. coli* (links), native Tyrosyl-tRNA von *Methanocaldococcus jannaschii* (mitte), für die Amber-Stop Codon Suppression optimierte tRNA aus *Methanocaldococcus jannaschii* (rechts) (Krahn *et al.*, 2020).

Das Fehlen einer Anticodon-Bindedomäne in der *MjTyrRS* gab Grund zur Annahme, dass sich eine Mutation des Anticodons von GUA zu CUA nicht auf die molekulare Erkennung der Suppressor-tRNA durch die *MjTyrRS* auswirken sollte. Allerdings interagiert das Anticodon mit Asp²⁸⁶ in der C-Terminalen Domäne des anderen Monomers, weshalb diese Aminosäure zu Arg²⁸⁶ mutiert wurde (Kobayashi *et al.*, 2003; Wang *et al.*, 2000). Das natürliche Substrat der *MjTyrRS*, Tyr, wird durch Wasserstoffbrücken zu den Aminosäuren Gln¹⁷³, Gln⁷⁵, Tyr¹⁵¹, Tyr³², Asp¹⁵⁸ erkannt. Dabei sind Tyr³² und Asp¹⁵⁸ für die Substratspezifität dieser Aminoacyl-tRNA Synthetase entscheidend (Brick & Blow, 1987).

Für die Generierung einer Aminoacyl-tRNA Synthetase, ausgehend von der *MjTyrRS*, die eine nicht-natürliche Aminosäure spezifisch erkennen soll, stehen daher die sechs genannten Aminosäuren im Fokus. Ausgehend von der strukturellen Information der *MjTyrRS* (Kobayashi *et al.*, 2003) wurden Gen-Bibliotheken des aktiven Zentrums der Aminoacyl-tRNA Synthetase mit einer hohen Diversität von mehr als 10⁹ Varianten erzeugt (Wang *et al.*, 2001; Wang *et al.*, 2006; Baumann *et al.*, 2019a). Das erfolgt zumeist mittels fokussierter Sättigungsmutagenese *in vitro* und wird anschließend in Plasmide kloniert. Diese auf den Plasmiden kodierte DNA-Bibliothek wird in die Wirtszelle transformiert, wobei die maximale Transformationseffizienz die Komplexität der Bibliothek festlegt. Die Qualität der Gen-Bibliotheken ist daher mit der Transformationseffizienz verflochten und ist der Grund, warum in den meisten Fällen *E. coli* Zellen, die für eine hohe Kompetenz nach vorangegangener Präparation bekannt sind, verwendet werden. *In vivo* können auch Gen-Bibliotheken über Mutator-Stämme wie *E. coli* XL1-red (Anderson *et al.*, 2004) und Mutator-Plasmide (Badran & Liu, 2015) generiert werden.

Nach der Generierung der Gen-Bibliothek für die Aminoacyl-tRNA Synthetase müssen jene Varianten selektiert werden welche die nicht-natürliche Aminosäure erkennen, diese adenlieren und auf die Suppressor-tRNA laden können (Positiv-Selektion). Diese gerichtete Selektion wird über ein Reporterprotein durchgeführt, welches in der DNA-Sequenz ein bis

sieben Amber-Stop Codons trägt. Als Reporterprotein werden Proteine verwendet, die eine Resistenz gegenüber Antibiotika vermitteln (Wang & Schultz, 2001; Liu & Schultz, 1999) oder ein Fluoreszenzsignal aufgrund der Suppression des Amber-Stop Codons erzeugen (Kuhn *et al.*, 2010). Der Vorteil bei der Selektion aus der Aminoacyl-tRNA Synthetase-Bibliothek über ein Fluoreszenzprotein ist, dass dies mit einem Durchflusszytometer durchgeführt und dabei der Selektionsdruck einfach angepasst werden kann. Zusätzlich wird weniger der oftmals sehr teuren nicht-natürlichen Aminosäure benötigt.

Die Selektion auf die Fähigkeit, die Suppressor-tRNA mit einer Aminosäure zu beladen und das Amber-Stop Codon zu supprimieren, bringt jedoch nicht ausschließlich Aminoacyl-tRNA Synthetasen hervor, die selektiv die nicht-natürliche Aminosäure erkennt. Darunter sind nach dieser Positiv-Selektion auch Varianten, die eine natürliche Aminosäure erkennen und zu deren ribosomalen Einbau an einem Amber-Stop Codon in die Polypeptidkette führen. Um diese Varianten abzureichern, muss die Aminoacyl-tRNA Synthetase-Bibliothek auch einer Negativ-Selektion unterworfen werden, dass in Abwesenheit der nicht-natürlichen Aminosäure über die Expression eines Reporterproteins mit einem Amber-Stop Codon in dessen kodierender Sequenz erfolgt. Zu diesem Zweck wurden häufig die Proteine Barnase (Wang & Schultz, 2001), CcdB Toxin (Umehara *et al.*, 2012) oder TolC (Amiram *et al.*, 2015) verwendet, die allerdings auf einem zweiten Plasmid kodiert waren, was die Umklonierung der Bibliothek erforderte. Die molekulare Erkennung einer natürlichen Aminosäure durch eine Aminoacyl-tRNA Synthetase-Variante führt zur Suppression des Amber-Stop Codons und resultiert in einem für die *E. coli* Zelle toxischem Protein. Für die Negativ-Selektion kann jedoch auch erneut das zuvor für die Positiv-Selektion benutzte Fluoreszenzprotein verwendet werden, da die Auswahl der zu selektierenden Zellen am Durchflusszytometer auch hinsichtlich ausbleibender Fluoreszenz möglich ist. Ein Umklonieren der Aminoacyl-tRNA Synthetase-Bibliothek ist durch die Verwendung dieses *One-plasmid* Systems nicht nötig (Kuhn *et al.*, 2010). Da, eine einfache Positiv-Selektion gefolgt von einer einfachen Negativ-Selektion bei Bibliotheken mit einer hohen Diversität nicht direkt zum Erfolg führt, werden diese beiden Selektionsstrategien abwechselnd durchgeführt, bis eine geringe Anzahl an geeigneten Aminoacyl-tRNA Synthetasen ausgewählt wurde.

1.4 Strategien zur Steigerung an Ausbeute an Proteinen mit nicht-natürlichen Aminosäuren

Da bei der Amber-Stop Codon-Technologie dem besagten Codon (UAG) eine Aminosäure, in diesem Fall eine nicht-natürlichen Aminosäure, neu zugewiesen wird, darf man dessen primäre und auch noch persistente Funktion dieser nicht außer Acht lassen, die Termination der Translation durch den Release Faktor 1 (Rf-1). Dieser konkurriert mit der Suppressor-tRNA bei der Anwendung der Amber-Stop Codon-Technologie um das Amber-Stop Codon

(Abbildung 5). Das führt zu einer verminderten Ausbeute an rekombinantem Protein, an dem am Amber-Stop Codon eine nicht-natürliche Aminosäure eingebaut werden soll, da die ribosomale Proteinbiosynthese beim überwiegenden Teil der mRNAs vorzeitig terminiert wird. Die dadurch verminderte Ausbeute kann man anhand von natürliche Suppressor-tRNAs (Eggertsson & Söll, 1988) abschätzen, wie zum Beispiel SupE aus dem *E. coli* Suppressorstamm XL-1 Blue, die abhängig von der mRNA Sequenz zwischen 12 – 26 % der Amber-Stop Codons supprimiert (Nene & Glass, 1981)

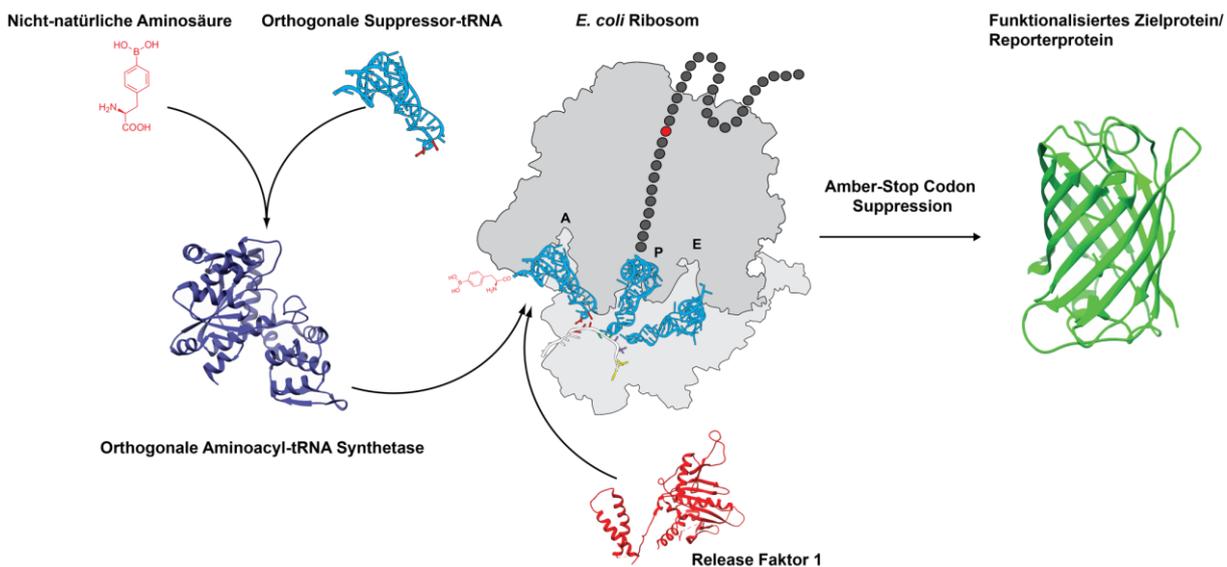


Abbildung 5: Ribosomaler Einbau einer nicht-natürlichen Aminosäure in ein Zielprotein mit Hilfe der Amber-Stop Codon-Technologie. Die Suppressor-tRNA konkurriert mit dem Rf-1 um das Amber-Stop Codon auf der mRNA.

Daher gab es Versuche diese Konkurrenz durch die Inaktivierung des Rf-1 zu beenden. Eine Deletion des korrespondierenden Gens (*pfrA*) schien theoretisch möglich, da der Rf-1 zwar neben dem am seltensten Amber-Stop Codon auch das am häufigsten vorkommende Ochre-Stop Codon terminiert, wohingegen letzteres zusammen mit dem Opal-Stop Codon auch vom Rf-2 terminiert wird. Jedoch erwies sich die Deletion der *prfA*-Gens in einem *E. coli* K12 Stamm als letal (Gerdes *et al.*, 2003; Rydén & Isaksson, 1984). Es war also zu vermuten, dass sich unter den dadurch nicht mehr korrekt terminierten Operons auch essenzielle Gene befinden. Daher wurden die Amber-Stop Codons der essentiellen Gene gegen Ochre-Stop Codons ausgetauscht und der Rf-1 deletiert (Gerdes *et al.*, 2003; Mukai *et al.*, 2010). In diesen Experimenten wurden aber nur *E. coli* K12 Derivate verwendet. Erst eine erfolgreiche Deletion des *prfA*-Gens in einem *E. coli* B-Stamm ermöglichte das wahre Problem zu erkennen. Der Rf-2 in *E. coli* K12 Derivaten trägt eine Mutation (A246T) und reduziert die Terminationseffizienz um den Faktor 5 (Johnson *et al.*, 2012; Johnson *et al.*, 2011), das aktuell als Grund für die letalen Folgen der *prfA*-Deletion in diesen *E. coli* Stämmen angesehen wird.

Der wohl prominenteste Ansatz zur Steigerung der Ausbeute an Proteinen mit nicht-natürlichen Aminosäuren der Forschungsgruppe um George Church tauschte daher die 321 Amber-Stop Codons in einem *E. coli* K12 Derivat gegen Ochre-Stop Codons und mutierte zusätzlich den endogene Rf-2 (T246A) um die reduzierte Terminationseffizienz in diesem Stamm auf das Level eines *E. coli* B Stammes zu steigern. Allerdings ist mit diesen Stämmen (C321.ΔA. und dessen Derivate) eine Kultivierung bei 30 °C empfohlen und zeigen eine um mindestens 25 % reduzierte Wachstumsrate (Lajoie *et al.*, 2013; Kuznetsov *et al.*, 2017).

Ein andersartiger Ansatz, den die Firma Sutro Biopharma (South San Francisco, USA) zur Steigerung der Ausbeute an rekombinanten nicht-natürlichen Proteinen verfolgt, beendet die Konkurrenz zwischen Rf-1 und der Suppressor-tRNA um das Amber-Stop Codon im Fall der präparativen *in vitro* Translation (Yin *et al.*, 2017). Dafür wurde eine geeignete Position am Rf-1 ausgewählt und zu einer Signalsequenz für die OmpT Protease mutiert. Während der Zellyse für die Herstellung des Translationsextrakts wird die Aspartyl-Protease OmpT, die sich normalerweise auf der äußeren Membran der *E. coli* B Zellen befindet, um Fremdproteine zu verdauen, freigesetzt. So kommt die Protease mit dem sich ebenfalls im Zellysat befindenden und Rf-1 in Kontakt und dieser wird an dem OmpT-Erkennungsmotif proteolytisch verdaut.

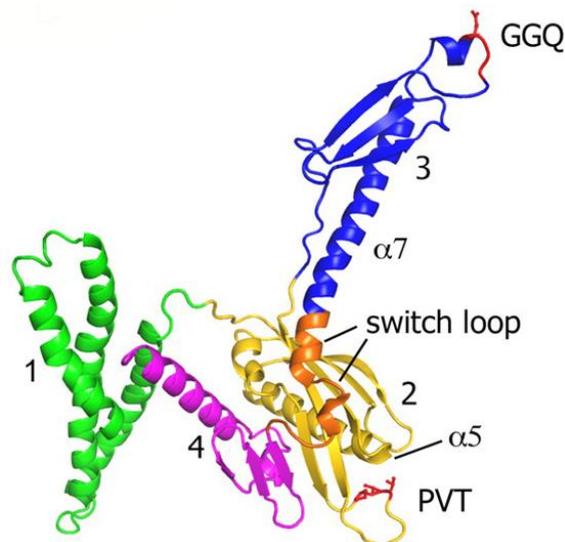


Abbildung 6: Darstellung des Rf-1 als Cartoon-Modell. Zu sehen sind die Domäne 1 (grün), die für die Erkennung des Amber-Stop Codons relevante Domäne 2 mit dem konservierten PVT Motif (gelb), die Domäne 3 mit dem konservierten GGQ Motif, das die chemische Bindung zwischen der Polypeptidkette und der tRNA im Peptidyl-Transferase-Zentrum des Ribosoms hydrolysiert (blau), die Domäne 4 (magenta) und der *Switch Loop* (orange) (Korostelev *et al.*, 2010).

Die Auswahl der geeigneten Position für die Mutation zu dem OmpT-Erkennungsmotif fiel auf den *Switch Loop* des Rf-1 und basiert auf der Struktur sowie auf der Funktion desselben. Der Rf-1, der in zwei Konformationen in der Zelle vorkommt, besteht aus 4 Domänen und dem *Switch Loop*. Die freie, geschlossene Konformation liegt im Cytoplasma vor und bindet mit dem PTV-Motif der Domäne 2 an das Amber- oder Ochre-Stop Codon der mRNA in der A-Position des Ribosoms (Laurberg *et al.*, 2008). Die Interaktion von Stop Codon und Rf-1 induziert eine

Konformationsänderung des Rf-1 zu der gebundenen, offenen Form und bringt das GGQ-Motif der Domäne 3 in das Peptidyl-Transferase-Zentrum des Ribosoms (P-Position). Die A-Position wird währenddessen von den Domänen 1 und 4 blockiert. Verantwortlich für diese Konformationsänderung ist der *Switch Loop*, der die Domänen 3 und 4 miteinander verbindet. Obwohl die 17 Aminosäuren der Primärsequenz des Switch Loops in Prokaryoten sehr divers sind, sind die zwei der Aminosäuren Arg³⁰³ und Ser³⁰⁴ konserviert (Korostelev *et al.*, 2010).

1.5 Zielsetzung der Arbeit

Im Rahmen dieser Arbeit sollte die Erweiterung des genetischen Codes genutzt werden, um einen synthetischen Kohlenhydrat-Rezeptor (Borocalin) in *E. coli* zu produzieren. Diese Borocalin-Variante sollte zur strukturellen Aufklärung der Protein/Kohlenhydrat-Interaktion, vermittelt durch die Boronsäure, kristallisiert werden und als Basis die Modifizierung der Kohlenhydrat-Bindetasche dienen. Zudem sollte eine spezifische Aminoacyl-tRNA Synthetase, ausgehend von der *MjTyrRS*, für die molekulare Erkennung einer fluoreszierenden nicht-natürlichen Aminosäure selektiert werden. Grundlage dafür bot sowohl eine bereits synthetisierte als auch eine eigenhändig und in Anlehnung an homologe Aminoacyl-tRNA Synthetasen mit ähnlicher Substratpräferenz generierte Gen-Bibliothek. Abschließend sollten speziell für die Amber-Stop Codon Suppression zugeschnittene *E. coli* Stämme mit steuerbarer Rf-1 Aktivität generiert werden, um die Ausbeute an nicht-natürlichem Protein bei der rekombinanten Expression zu steigern.

2 Material und Methoden

2.1 Material

2.1.1 Bakterienstämme und Plasmide

Escherichia coli-Stämme:

BL21	F ⁻ , <i>ompT</i> , <i>lon</i> , <i>hsdS_B</i> (r _B ⁻ m _B ⁻), <i>gal</i> , <i>dcm</i> (Studier & Moffatt, 1986)
XL1-blue	<i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>hsdR17</i> (r _k ⁻ , m _k ⁺), <i>supE44</i> , <i>relA1</i> , Δ(<i>lacZYA-argF</i>)U169 [F ['] : <i>proAB</i> , <i>lacI^q</i> ZΔM15, Tn10(<i>Tet^r</i>)] (Bullock, 1987)
NEB Express	<i>fhuA2</i> [<i>lon</i>], <i>ompT</i> , <i>gal</i> , <i>sulA11</i> , R(<i>mcr-73::miniTn10--Tet^S</i>) ₂ [<i>dcm</i>], R(<i>zgb-210::Tn10--Tet^S</i>), <i>endA1</i> Δ(<i>mcrC-mrr</i>)114:: <i>IS10</i> (New England Biolabs, Frankfurt a. M.)
MG1655	F ⁻ , λ ⁻ , <i>ilvG⁻</i> , <i>rfb-50</i> , <i>rph-1</i> (Blattner et al., 1997)

Plasmide:

pSB4.6c2	Dr. Selvakumar Edwardraja, Lehrstuhl für Biologische Chemie, 2016
pSB4.6c14 (NFW)	Dr. Selvakumar Edwardraja, Lehrstuhl für Biologische Chemie, 2016
pSB6.6c14 (R2C2)	Dr. Selvakumar Edwardraja, Lehrstuhl für Biologische Chemie, 2016
pSB6.6c14 (R2C2.1)	Dr. Selvakumar Edwardraja, Lehrstuhl für Biologische Chemie, 2016
pSB6.6c14 (R2C2.7)	Dr. Selvakumar Edwardraja, Lehrstuhl für Biologische Chemie, 2016
pSB15.CafRS#30d53	Andreas Reichert, Lehrstuhl für Biologische Chemie, 2018
pSB16.23c2	Dr. Peter Mayrhofer, Lehrstuhl für Biologische Chemie, 2020
pSB15.100d2	Dr. Peter Mayrhofer, Lehrstuhl für Biologische Chemie, 2020
pCas	Sheng Yang, Addgene # 62225 (Jiang et al., 2015)

pTargetF	Sheng Yang, Addgene # 62226 (Jiang <i>et al.</i> , 2015)
pASK75	Labor Prof. Skerra, TU München (Skerra, 1994)
pKD46	Dr. Martin Schlapschy, Lehrstuhl für Biologische Chemie, 2019 (Datsenko & Wanner, 2000)
pRK793	Uli Binder, XL-Protein GmbH, Freising, 2019 (Kapust <i>et al.</i> , 2001)
pUC18	Ina Theobald, Lehrstuhl für Biologische Chemie, 2019 (Yanisch-Perron <i>et al.</i> , 1985)
13ABRFYP-1415337-P1A5	XL-Protein GmbH, Freising, 2019

2.1.2 Gensynthesen und Oligodesoxynukleotide

Gensynthesen

Die Gene für die Aminoacyl-tRNA Synthetasen AzAlaRS, PhotoDopaRS und HqAlaRS wurden mit optimierten Codons für in *E. coli* von der Firma von Thermo Fisher Scientific, Regensburg, als doppelsträngige DNA-Fragmente synthetisiert. Dabei wurde darauf geachtet, dass die Codon-Nutzung den zuvor synthetisierten Gen für die BpaRS (Edwardraja *et al.*, 2017) entsprach und am 5'- bzw. 3'-Ende die Primer ep1 und ep2 zur Amplifikation hybridisieren.

AzAlaRS (PDB ID: 5L7P):

```
5' -CTGAAATTATCAGCGAGGAAGAGCTGCGTGAGGTGCTGAAGAAAGATGAGAAATCGGCGGGGATC
GGCTTTGAACCGAGCGGTAAAATCCATCTGGGCCACTATCTGCAAAATCAAGAAAATGATCGATCTGCA
AAATGCCGGCTTCGACATTATTATCATCCTGGCGGACCTGGCTGCATATTTAAATCAGAAAGGCGAAC
TGGATGAAATCCGTAAAATTTGGTGACTATAACAAGAAAGTGTTTCGAGGCCATGGGTCTGAAAGCCAAA
TATGTGTATGGCTCCGAGTTTGCTCTGGACAAAGACTATAACCCTGAACGTGTACCGGTTGGCACTGAA
AACCACGCTGAAACGTGCTCGTCGTAGCATGGAACCTGATTGCCCGTGAGGACGAAAATCCGAAAGTTG
CTGAGGTGATTTATCCGATTATGCAAGTGAACGCCGCACATTATATGGGCGTCGATGTTAATGTCCGT
GGCATGGAGCAGCGTAAAATCCACATGCTGCAGCGTGAACCTGCTGCCTAAGAAAGTCGTGTGTATCCA
CAATCCGGTACTGACTGGTCTGGACGGTGAAGGCAAAAATGAGCAGCTCTAAAGGTAACCTTTATTGCCG
TGGATGATTCTCCAGAAGAGATTAGGGCTAAG-3'
```

PhotoDopaRS (PDB ID: 5NSF):

```
5' -CTGAAATTATCAGCGAGGAAGAGCTGCGTGAGGTGCTGAAGAAAGATGAGAAATCGGCGGGGATC
GGCTTTGAACCGAGCGGTAAAATCCATCTGGGCCACTATCTGCAAAATCAAGAAAATGATCGATCTGCA
AAATGCCGGCTTCGACATTATTATCTGGCTGGCGGACCTGGGGGCATATTTAAATCAGAAAGGCGAAC
TGGATGAAATCCGTAAAATTTGGTGACTATAACAAGAAAGTGTTTCGAGGCCATGGGTCTGAAAGCCAAA
TATGTGTATGGCTCCGAGCACAACCTGGACAAAGACTATAACCCTGAACGTGTACCGGTTGGCACTGAA
AACCACGCTGAAACGTGCTCGTCGTAGCATGGAACCTGATTGCCCGTGAGGACGAAAATCCGAAAGTTG
```

CTGAGGTGATTTATCCGATTATGCAAGTGAACGCTATTCATTATAACGGCGTCGATGTTGCGGTCCGGT
GGCATGGAGCAGCGTAAAATCCACATGCTGGCCCCGTGAACGCTGCCTAAGAAAGTCGTGTGTATCCA
CAATCCGGTACTGACTGGTCTGGACGGTGAAGGCAAAAATGAGCAGCTCTAAAGGTAAC TTTATTGCCG
TGGATGATTCTCCAGAAGAGATTAGGGCTAAG-3'

HqAlaRS (Liu *et al.*, 2013):

5'-CTGAAATTATCAGCGAGGAAGAGCTGCGTGAGGTGCTGAAGAAAGATGAGAAATCGGCGCATATC
GGCTTTGAACCGAGCGGTAAAATCCATCTGGGCCACTATCTGCAAATCAAGAAAATGATCGATCTGCA
AAATGCCGGCTTCGACATTGTAATCCATCTGGCGGACCTGGGAGCATATTTAAATCAGAAAGGCGAAC
TGGATGAAATCCGTAAAATTTGGTACTATAACAAGAAAGTGTTCGAGGCCATGGGTCTGAAAGCCAAA
TATGTGTATGGCTCCGAGCGGGTCTGGACAAAGACTATACCCTGAACGTGTACCGGTTGGCACTGAA
AACCACGCTGAAACGTGCTCGTCGTAGCATGGAAC T GATTGCCCGTGAGGACGAAAATCCGAAAGTTG
CTGAGGTGATTTATCCGATTATGCAAGTGAACAATATTCATTATGACGGCGGTGATGTTGCGGTCCGGT
GGCATGGAGCAGCGTAAAATCCACATGCTGGCCCCGTGAACGCTGCCTAAGAAAGTCGTGTGTATCCA
CAATCCGGTACTGACTGGTCTGGACGGTGAAGGCAAAAATGAGCAGCTCTAAAGGTAAC TTTATTGCCG
TGGATGATTCTCCAGAAGAGATTAGGGCTAAG-3'

Primer zur Sequenzierung und Amplifizierung von Genen

CS_seq1	5'-CTGAAATTATCAGCGAGGAAGAGCTGCG-3'
CS_seq2	5'-GAACGTGGCGAGAAAGGAAG-3'
CS_seq3	5'-GGCGCACATTGTGCGAC-3'
CS_seq5	5'-GGGCGTGAATTTAACGCGG-3'
CS_seq6	5'-CTCATCGCAGTACTGTTG-3'
CS_seq7	5'-AGCGAGTCAGTGAGCGAG-3'
CS_seq8	5'-GCAGGATCTGGCATGGAA-3'
CS_seq9	5'-CCTGCGGGAAGTTGTAAG-3'
CS_seq10	5'-GTAACGCCTTCCTCGAAG-3'
CS_seq13	5'-TGTA AACGACGGCCAGT-3'
CS_seq14	5'-ACACTTTATGCTTCCGGC-3'
CS_seq15	5'-GATGAAGCCCTGAAAGAC-3'
CS_seq16	5'-CACATGGTCCTTCTTGAG-3'

CS_seq17	5' -CCGCGTTAAATTCACGCCC-3'
D20	5' -CCACTCCCTATCAGTGAT-3'
Seq 2	5' -CCTTCAACAACCTCCAGCTG-3'
MW6	5' -CGGATAACAATTTACACAG-3'
Primer 1	5' -CGCAGTAGCGGTAAACG-3'
ep1	5' -CTGAAATTATCAGCGAGGAAGAG-3'
ep2	5' -CCTAGCCCTAATCTCTTCTGGAG-3'
rfCS_5	5' -TCGTTAGAATTCGGTAGAAATCATGAGCGCCAGC-3'
rfCS_6	5' -TCGCCCTGGAAGTAGAGATTTTCACGGGTAGACGCTTCGGC-3'
rfCS_7	5' -AAATCTCTACTTCCAGGGCGATCGCAGCGACCGTA-3'
rfCS_8	5' -GCAACGACCAGAATTCTCGCAC-3'
rfCS_19	5' -TTATTAGAATTCGGCGAATGCTTATCTTTATAGCG-3'
rfCS_20	5' -TCATAGGCGTAAATGCACCCTG-3'
rfCS_21	5' -CAGGGTGCATTTACGCCTATGATTGAACAGGATGGCCTG-3'
rfCS_22	5' -CACGTAACCAGTGTGATATTCCATTTAGAAAAATTCATCCAGCAG ACGAT-3'
rfCS_23	5' -ATGGAATATCAACACTGGTTACGTG-3'
rfCS_24.2	5' -AAAATTGAATTCCTAAACCAGTCGCTTTGCAGAATG-3'
rfCS_48	5' -AATAATAAGCTTAACGAGGGCAAAAAATGAAATCGGACTCAGAAG TCAATC-3'
rfCS_49	5' -GATTTTCAGAAGCGCCACCAATCTGTTCTCTG-3'
rfCS_50	5' -GTGGCGCTTCTGAAAATCTTTATTTTCAAGGTCATCATC-3'
rfCS_51	5' -CAGGTCAAGCTTTTATTAATTCATGAGTTGAGTCGCTTC-3'

Primer für die QuikChange-Mutagenese

CS19_36Bpa R2C2. G72A_Y132W	5' -CTCAAAACAGGGAGTGGTTCAACATCACCCCTC-3'
CS20_36Bpa R2C2. G72A_Y132W	5' -GAGGGTGATGTTGAACCACTCCCTGTTTTGAG-3'
CS22_36Bpa NFW G72A_F52W	5' -CAAAAGATGTGGGCCACCATCTATG-3'

CS23_36Bpa NFW G72A_F52W	5' -CATAGATGGTGGCCACATCTTTTG-3'
CS24_36Bpa NFW N134A	5' -GGGAGTACTTCGCAATCACCTCTACG-3'
CS25_36Bpa NFW N134A	5' -CGTAGAGGGTGATTGCGAAGTACTCCC-3'
CS26_36Bpa NFW W79Y	5' -GAAGTGTGACTACTACATCAGGACTTTTGTTCAGG-3'
CS27_36Bpa NFW W79Y	5' -CCTGGAACAAAAGTCCTGATGTAGTAGTCACACTTC-3'
rfCS_30	5' -CATCATCATCATGGAGAATCCTTGTTTAAGGGGCC GCG-3'
rfCS_31	5' -CGCGGCCCTTAAACAAGGATTCTCCATGATGATG ATG-3'
rfCS_37	5' -GGGCTTAAGCTTAACGAGGGCAAAAAATGAAACATCA TCATCATCATCATCATGG-3'
rfCS_38	5' -TTGGTTAAGCTTTTATTAATTCATGAGTTGAGTCGC-3'

Oligodesoxynukleotide für die Hybridisierung zu DNA-Fragmenten

LibCS_9	5' -CGAGGAAGAGCTTTATTGCCGTGGACGTGATTC-3'
LibCS_10	5' -GGAGAATCACGTCCACGGCAATAAAGCTCTTCC-3'
rfCS_54	5' -GTGGCTGGGGTACGAGGATTCGAACCTCGGAATGCCGG-3'
rfCS_55	5' -TCCGAGGTTCGAATCCTCGTACCCCAGCCACTGCA-3'
rfCS_56	5' -AATTAGAATCCGGTGCCTTACCGCTTGCGATACCCCAG-3'
rfCS_57	5' -AATTCTGGGGTATCGCCAAGCGGTAAGGCACCGGATTCTAATTCCG GCAT-3'

Alle Oligodesoxynukleotide wurden von Eurofins Genomics, Ebersberg, in RP-HPLC-gereinigter Form bezogen.

2.1.3 Enzyme und sonstige Proteine

<i>Pfu</i> Ultra II Fusion HS DNA-Polymerase mit dazugehörigem Herstellerpuffer	Agilent Technologies, Santa Clara, CA, USA
Q5 High-Fidelity DNA-Polymerase mit dazugehörigem Herstellerpuffer	New England Biolabs, Frankfurt a. M.
Hot Strat Taq DNA-Polymerase mit dazugehörigem Herstellerpuffer	New England Biolabs, Frankfurt a. M.
DNaseI	New England Biolabs, Frankfurt a. M.

Benzonase®, Purity stage II (25U/μl)	Merk, Darmstadt
Restriktionsendonukleasen mit dazugehörigem Herstellerpuffer	New England Biolabs, Frankfurt a. M. Thermo Fisher Scientific, St. Leon-Rot
T4 DNA-Ligase (1 U/μl und 5 U/μl) mit dazugehörigem Herstellerpuffer	Fermentas, St. Leon-Rot
T4-Polynukleotid-Kinase mit dazugehörigem Herstellerpuffer	New England Biolabs, Frankfurt a. M.
FastAP™ Thermosensitive Alkalische Phosphatase	Thermo Fisher Scientific, St. Leon-Rot

2.1.4 Chemikalien

1,4-Dithiothreitol (DTT)	Thermo Fisher Scientific, St. Leon-Rot
2'-Desoxyadenosin 5'-triphosphat (dATP)	Thermo Fisher Scientific, St. Leon-Rot
2'-Desoxycytidin 5'-triphosphat (dCTP)	Thermo Fisher Scientific, St. Leon-Rot
2'-Desoxyguanosin 5'-triphosphat (dGTP)	Thermo Fisher Scientific, St. Leon-Rot
2'-Desoxythymidin 5'-triphosphat (dTTP)	Thermo Fisher Scientific, St. Leon-Rot
2'-(p-Hydroxyphenylazo)-benzoesäure (HABA), 98 %	Alfa Aesar, Heysham, UK
2-(N-Morpholino)ethansulfonsäure (MES)	Carl Roth, Karlsruhe
4'-Hydroxyazobenzol-2-carbonsäure (HABA)	Thermo Fisher Scientific, St. Leon-Rot
Acetonitril, LC-MS grade (ACN)	Sigma-Aldrich Chemie, Taufkirchen
Acrylamid-Stammlösung 30 % (w/v) mit 0,8 % (v/v) Bisacrylamid	SERVA, Heidelberg
Adenosin-5'-triphosphat (ATP)	AppliChem, Darmstadt
Agarose für die Gelelektrophorese	Bio&SELL, Feucht
Agarose (<i>Low Melting Point</i>)	Invitrogen, Darmstadt
Ammoniaklösung, 25 % (v/v) p.a.	Carl Roth, Karlsruhe
Ammoniumacetat, p.a.	Sigma-Aldrich Chemie, Taufkirchen
Ammoniumbicarbonat	Carl Roth, Karlsruhe
Ammoniumchlorid	AppliChem, Darmstadt
Ammoniumperoxodisulfat (APS)	Carl Roth, Karlsruhe
Ampicillin (Amp)	Carl Roth, Karlsruhe
Anhydrotetracyclin (aTc)	Acros Organics, Geel, BE
Bacto Agar	Difco Laboratories, Detroit, MI, USA
Bacto Trypton	Difco Laboratories, Detroit, MI, USA

Bacto Hefeextrakt	Difco Laboratories, Detroit, MI, USA
Bromphenolblau	AppliChem, Darmstadt
Calciumacetat, Hydrat	Sigma-Aldrich Chemie, Taufkirchen
Calciumchlorid, Dihydrat	AppliChem, Darmstadt
Cetyltrimethylammoniumbromid	Carl Roth, Karlsruhe
Chloramphenicol	Carl Roth, Karlsruhe
Chloroform	Merk, Darmstadt
Coomassie Brilliantblau R-250	AppliChem, Darmstadt
D-Desthiobiotin	IBA, Göttingen
Dinatriumhydrogenphosphat	AppliChem, Darmstadt
Dikaliumhydrogenphosphat	AppliChem, Darmstadt
Essigsäure, 99,9 %, p.a.	Carl Roth, Karlsruhe
Ethanol, p.a.	Sigma-Aldrich Chemie, Taufkirchen
Ethanol, technisch	CLN Laborgroßhandel, Niederhummel
Ethidiumbromid	AppliChem, Darmstadt
Ethylendiamintetraessigsäure (EDTA)	AppliChem, Darmstadt
Glycerin, 99 %	Carl Roth, Karlsruhe
Glycin	AppliChem, Darmstadt
Imidazol	Fluka Chemie, Buchs, Schweiz
Isoamylalkohol	Merk, Darmstadt
Isopropanol, p.a.	Carl Roth, Karlsruhe
Isopropanol, technisch	CLN Laborgroßhandel, Niederhummel
Kanamycin A	Carl Roth, Karlsruhe
Kaliumdihydrogenphosphat	AppliChem, Darmstadt
β -Mercaptoethanol	Sigma-Aldrich Chemie, Taufkirchen
Magnesiumchlorid, Hexahydrat	AppliChem, Darmstadt
Magnesiumsulfat, Heptahydrat	Carl Roth, Karlsruhe
N-Cyclohexyl-2-aminoethanesulfonic acid (CHES)	AppliChem, Darmstadt
N,N,N',N'-Tetramethylethylendiamin (TEMED)	Carl Roth, Karlsruhe
N-Acetyl-L-tryptophanamid	Sigma-Aldrich Chemie, Taufkirchen
Natriumazid, ≥ 99 %, p.a.	Carl Roth, Karlsruhe
Natriumchlorid	Carl Roth, Karlsruhe
Natriumdodecylsulfat (SDS)	Carl Roth, Karlsruhe

Natriumhydroxid, Plätzchen, p.a.	Carl Roth, Karlsruhe
Puffer-Standardlösungen für pH-Meter	Carl Roth, Karlsruhe
Roti [®] -Aqua-Phenol	Carl Roth, Karlsruhe
Salzsäure, 37 %, p.a.	Carl Roth, Karlsruhe
Spectinomycin Dihydrochlorid Pentahydrat	AppliChem, Darmstadt
Stickstoff N50, CO-frei	Linde, Pullach
Thiamin-hydrochlorid	Sigma-Aldrich Chemie, Taufkirchen
Tris(hydroxymethyl)aminomethan (Tris)	AppliChem, Darmstadt

2.1.5 Zucker und Zuckerderivate

4-Nitrocatechol	Sigma-Aldrich Chemie, Taufkirchen
D-Galactose	AppliChem, Darmstadt
D-Glucose	Carl Roth, Karlsruhe
D-Mannose	Fluka, St Gallen (jetzt Sigma-Aldrich Chemie, Taufkirchen)
L-Fucose	SERVA Electrophoresis, Heidelberg
Lewis Y	Carbosynth, UK
N-Acetylglucosamin	Carbosynth, UK
Sorbit	Sigma-Aldrich Chemie, Taufkirchen

Für die Titration wurden die Zucker und Zuckerderivate in ddH₂O gelöst.

2.1.6 (Nicht-)natürliche Aminosäuren

Tyrosin	Fluka, St Gallen (jetzt Sigma-Aldrich Chemie, Taufkirchen)
D/L-4-Borono-L-Phenylalanin (Bpa)	AppliChem, Darmstadt
7-Hydroxycoumarinalanin (Hca) als Hydrochlorid	Dr. Christopher Graf, Lehrstuhl für Biologische Chemie, 2018
Borono-2,6-difluorphenyl-alanin (Bfa) als Hydrochlorid	Dr. Markus Anneser, Lehrstuhl für Biologische Chemie, 2019
Azulenalanin (AzAla) als Natriumsalz	Dr. Markus Anneser, Lehrstuhl für Biologische Chemie, 2020
L-4-Azidophenylalanin (Apa)	Dr. Markus Anneser, Lehrstuhl für Biologische Chemie, 2020
Tetrahydrofuran-Lysin (THF-K)	Dr. Markus Anneser, Lehrstuhl für Biologische Chemie, 2020

2.1.7 Standards und Kits

DNA- und Proteingrößenstandards

GeneRuler 1 kb DNA Ladder	Thermo Fisher Scientific, St. Leon-Rot
GeneRuler 100 bp DNA Ladder	Thermo Fisher Scientific, St. Leon-Rot
PageRuler Prestained Protein Ladder	Thermo Fisher Scientific, St. Leon-Rot
Unstained Protein MW Marker	Thermo Fisher Scientific St. Leon-Rot

Kits zur DNA-Isolierung

QIAprep Spin Plasmid Mini Kit	Quiagen, Hilden
QIAquick Gel Extraction Kit	Quiagen, Hilden
QIAquick PCR Purification Kit	Quiagen, Hilden
Wizard [®] SV Gel / PCR Clean-Up System	Promega, Mannheim

Kit zur DNA-Sequenzierung

Mix2Seq-Sequenzierungskit	Eurofins Genomics, Ebersberg
---------------------------	------------------------------

2.1.8 Geräte

Elektrophorese

Beleuchtungstisch T13	Biometra, Göttingen
HE33 Submarine (DNA-Auftrennung)	Hoefler, Holliston, MA, USA
Minigel-Twin (Protein-Auftrennung)	Biometra, Göttingen
Power Pack P25	Biometra, Göttingen
UV-Transilluminator TI 3 (254 nm)	Biometra, Göttingen

Inkubatoren

Friocell	MMM Medcenter Einrichtungen, Planegg
Incucell	MMM Medcenter Einrichtungen, Planegg
Innova 43 series	MMM Medcenter Einrichtungen, Planegg
Minitron	Infors, Bottmingen, Schweiz
Multitron	Infors, Bottmingen, Schweiz

Pipetten und Pipettierhilfe

Discovery Comfort P20, P200, P1000	Corning HTL, Warschau, Polen
Eppendorf Research Plus, 0,5-10 µl	Eppendorf, Hamburg
Pipetboy acu	Integra Bioscience, Fernwald

Proteinreinigung

ÄKTA Pure	GE Healthcare, Freiburg
ÄKTA Purifier	GE Healthcare, Freiburg
Fraktionssammler Frac900	GE Healthcare, Freiburg
Peristaltische Pumpe P-1	GE Healthcare, Freiburg
Signalschreiber Rec 112	GE Healthcare, Freiburg

Spektralphotometer

NanoDrop 2000	Thermo Fisher Scientific, Bonn
Spektrphotometer Ultraspec 2100 pro	GE Healthcare, Freiburg
Spektrphotometer Ultraspec 3000	GE Healthcare, Freiburg

Waagen

Feinwaage 125A	Precisa Gravimetrics, Dietikon Schweiz
Laborwaage XB22C	Precisa Gravimetrics, Dietikon Schweiz
Laborwaage PE620	Gebr. Bosch, Jungingen

Wasserfilter- und Destillationsanlagen

ddH ₂ O Anlage Typ 2304	Gesellschaft für Laborbedarf, Würzburg
Reinstwasser Anlage Synergy 185	Merck, Darmstadt

Zentrifugen/Rotoren

Eppendorf 5418R	Eppendorf, Hamburg
Sigma 4-16K (11118 und 11156)	Sigma, Osterode
Sorvall RC 12BP+	Thermo Fisher Scientific, Schwerte
Sorvall RC 6 Plus	Thermo Fisher Scientific, Schwerte
H-12000 Bioprocessing-Ausschwingrotor	Thermo Fisher Scientific, Schwerte
SLA-1500 Super-Lite Aluminium Festwinkelrotor	Thermo Fisher Scientific, Schwerte
SLA-3000 Super-Lite Aluminium Festwinkelrotor	Thermo Fisher Scientific, Schwerte
SLA 4X1L LEX	Thermo Fisher Scientific, Schwerte
SS-34 Festwinkelrotor	Thermo Fisher Scientific, Schwerte

Sonstige Geräte

Autoklav Varioklav Typ 500	H&P Labortechnik, Oberschleißheim
maXis Q-TOF Massenspektrometer mit ESI-Ionenquelle	Bruker Daltonics, Bremen
pH-Meter inolab pH720	WTW, Weilheim
Thermocycler Mastercycler gradient	Eppendorf, Hamburg
Thermocycler TPersonal	Biometra, Göttingen
Vortex Genie 2	Scientific Industries, Bohemia, NY, USA
Wasserbad Haake F6/C25	Thermo Haake, Karlsruhe
Ultraschall Homogenisator, Digital Sonifier 250	Branson Ultrasonics Corporation, Connecticut, US
FACSAria II	Becton Dickinson, New Jersey, USA
RVC 2-18 CDplus (Speed-Vac)	Christ, Osterode am Harz
Spektrofluorimeter LS 50 B	Perkin Elmer, Massachusetts, USA
MicroPulser Electroporator	BioRad, Kalifornien, USA

Weitere zur Grundausstattung eines biotechnologischen Labors gehörende Geräte sind hier nicht explizit genannt.

2.1.9 VerbrauchsmaterialeienChromatographie-Säulen und Material

Resource S, 1 bzw. 6 ml	GE Healthcare, Freiburg
<i>Strep</i> -Tactin Superflow	IBA, Göttingen
Superdex 200 HiLoad 16/60	GE Healthcare, Freiburg
Ni-Sepharose High Performance, 1 ml	Cytiva, Marlborough, Massachusetts, USA

Sonstiges Material

Accudrop Beads	Becton Dickinson, New Jersey, USA
Amicon Ultrafiltrationsröhrchen, 4/15 ml (MWCO: 10 kDa)	Merck, Darmstadt
Atmungsaktive Verschlussfolie für Mikrotiterplatten	Greiner Bio-One, Frickenhausen
Hefe-tRNA	Thermo Fisher Scientific, St. Leon-Rot
CS&T research beads	Becton Dickinson, New Jersey, USA
Quarz-Küvette Typ 110-QS mit Deckel (Schichtdicke: 0,1 cm)	Hellma, Müllheim

Dialysemembran Typ 27 (22 mm, MWCO: 14 kDa)	Precisa Gravimetrics, Dietikon, Schweiz
Einweg-Spritzen Omifix (2, 10, 30, 50 ml)	Braun, Melsungen,
Erlenmeyerkolben (50, 100, 250, 1000, 5000 ml)	Duran Group, Wertheim
Filtropur Spritzenfilter (0,22 bzw. 0,45 μm)	Sarstedt, Nürnbergrecht
Glaspipetten (5, 10, 25 ml)	Sarstedt, Nürnbergrecht
Kulturröhrchen, 13 ml	Sarstedt, Nürnbergrecht
Laborglasflaschen (25, 50, 100, 500, 1000, 2000 ml)	Duran Group, Wertheim
Membranfilter Microsart CN, 0,45 μm	Satorius, Göttingen
Membranfilter NC20, 0,2 μm	GE Healthcare, Freiburg
Parafilm	Pechiney Plastic Packaging, WI, USA
Petrischalen, rund und eckig	Sarstedt, Nürnbergrecht
Pipettenspitzen:	
Kristallklar (0,5 bis 10 μl)	Brand, Wertheim
Gelb (10-200 μl)	Sarstedt, Nürnbergrecht
Blau (100-1000 μl)	Gilson International, Limburg-Offheim
Reaktionsgefäße (0,2, 0,5, 1,5, 2,0 ml)	Eppendorf, Hamburg
Slide-A-Lyzer Mini Dialysis Unit (MWCO: 10 kDa, Vol.: 10-100 μl)	Thermo Fisher Scientific, St. Leon-Rot
Universal-Indikatorstäbchen pH-Fix, pH 0-14	Carl Roth, Karlsruhe
UV-VIS QS (1 cm)	Hellma, Müllheim
VIS Halbküvetten, Plastik	Sarstedt, Nürnbergrecht
Vivaspin 500 (MWCO: 10 kDa)	Sartorius, Göttingen
Zellschaber	Greiner Bio-One, Frickenhausen
Zentrifugenröhrchen (15, 50 ml)	Greiner, Frickenhausen
Zentrifugenbecher/-röhrchen (SLA1500, SS-34, H-12000)	Thermo Fisher Scientific, St. Leon-Rot
Falcon™ Rundboden-Polystyrolröhrchen	Thermo Fisher Scientific, St. Leon-Rot

2.1.10 Medien, Antibiotika und allgemeine Lösungen

Alle Medien sowie Lösungen für die Arbeit mit Bakterien, Proteinen und DNA wurden zur Sterilisation entweder für mind. 20 min bei 121 °C autoklaviert oder sterilfiltriert und die Chromatographiepuffer zusätzlich für mind. 30 min entgast. Sämtliche Lösungen mit Ausnahme der Kulturmedien, für die entionisiertes Wasser verwendet wurde, wurden mit

Reinstwasser (Typ I, 18,2 M Ω ·cm) hergestellt. Die verwendeten Kulturplatten wurden bei 4 °C, und Lösungen, falls nicht anders angegeben, bei Raumtemperatur aufbewahrt. Die Zugabe von Antibiotika erfolgte bei Festmedien vor dem Gießen nach Abkühlen des Mediums auf ca. 50 °C und bei Flüssigmedien erst unmittelbar vor der Verwendung.

Stammlösungen

Für die Herstellung von Selektivmedien bzw. zur Induktion der rekombinanten Genexpression wurden folgende Stammlösungen verwendet. Diese wurden bei -20 °C gelagert und vor Verwendung sterilfiltriert (0,45 μ m).

Ampicillin (Amp)	100 mg/ml in 50 % v/v Ethanol (1:1000 zum Medium)
Anhydrotetracyclin (aTc)	2 mg/ml in 70 % v/v Ethanol
Chloramphenicol (Cam)	30 mg/ml in 70 % v/v Ethanol (1:1000 zum Medium)
Kanamycin	50 mg/ml in 70 % v/v Ethanol (1:1000 zum Medium)
Spectinomycin	50 mg/ml in 70 % v/v Ethanol (1:1000 zum Medium)

LB-Medium

Bacto Trypton	10 g/l
Bacto Hefeextrakt	5 g/l
NaCl	5 g/l
Bacto Agar	15 g/l (nur für Kulturplatten)

Mit NaOH auf pH 7,5 eingestellt und autoklaviert

TB-Medium

TB-Puffer-Stammlösung (10x)

KH ₂ PO ₄	5 g/l
K ₂ HPO ₄	5 g/l

TB-Medium

Bacto Trypton	12 g/l
Bacto Hefeextrakt	24 g/l
Glycerin	4 ml/l

Die Zugabe der autoklavierten Pufferstammlösung erfolgte nach dem Abkühlen des autoklavierten Mediums im Verhältnis 1:10.

M9 Minimalmedium

Mineralsalzlösung:

31 mM Na ₂ HPO ₄ · 2 H ₂ O	5,51 g/l
19 mM KH ₂ PO ₄	2,58 g/l
25 mM NH ₄ Cl	1,33 g/l
5 mM Na ₃ -Citrat · 2 H ₂ O	0,47 g/l

Die Mineralsalzlösung stellt sich auf einen pH von 6,9 ein und wurde im Autoklaven sterilisiert.

Glucosestammlösung:

20 % w/v Glucose (separat autoklaviert)

MgSO₄-Stammlösung:1 M MgSO₄-Lösung (separat autoklaviert)

Thiamin-Stammlösung:

10 mg/ml Thiamin-Hydrochlorid (sterilfiltriert)

Medienbereitung:

45 ml	Mineralsalzlösung
5 ml	Glucose (20 %)
0,5 ml	MgSO ₄ (1 M)
50 µl	Thiamin (10 mg / ml)

SOC Medium

Bacto Trypton	20 g/l
Bacto Hefeextrakt	5 g/l
NaCl	0,5 g/l
KCl	2,5 mM

Der pH wurde vor dem Autoklavieren mit NaOH auf 7,5 eingestellt und anschließend je Liter folgende Volumina steriler Lösungen zugegeben:

1 M MgCl ₂	10 ml
1 M MgSO ₄	10 ml
20 % Glucose	20 ml

GYT Medium

Bacto Trypton	0,25 % (w/v)
Bacto Hefeextrakt	0,125 % (w/v)
Glycerin	10 % (v/v)

Puffer für die Agarose-Gelelektrophorese

TAE-Puffer:

Tris	40 mM
Essigsäure	20 mM
EDTA	1 mM

Auftragspuffer für Agarosegele (10x):

Bromphenolblau	1 mg/ml
Glycerin	50 % (v/v)
Xylencyanol	1 mg/ml

Puffer für die SDS-PAGE

SDS-Laufpuffer nach Fling and Gregerson (1986) (1x):

Glycin	190 mM
SDS	1 g/l
Tris	50 mM
pH 8,8 stellt sich ein	

Lower-Tris (4x):

SDS	4 g/l
Tris/HCl pH 8,85	3 M

Upper-Tris (4x):

SDS	4 g/l
Tris/HCl pH 6,8	0,5 M

Auftragspuffer (5x):

Bromphenolblau	0,25 mg/ml
Glycerin	25 % (v/v)
SDS	7,5 % (w/v)
β -Mercaptoethanol	12,5 % (v/v) (bei reduzierendem Auftragspuffer)

Färbelösung:

Coomassie Brilliantblau R-250	1 % (w/v)
Essigsäure	10 % (v/v)
Isopropanol	25 % (v/v)

Entfärbelösung I:

Essigsäure	10 % (v/v)
Isopropanol	25 (v/v)

Entfärbelösung II:

Essigsäure	10 % (v/v)
------------	------------

Puffer für die Anionenaustauschchromatographie

Laufpuffer:

MES/HCl (pH 6,0)	20 mM
------------------	-------

Elutionspuffer:

MES/HCl (pH 6,0)	20 mM
------------------	-------

NaCl	500 mM
------	--------

Puffer für die Metallchelate-Affinitätschromatographie

Laufpuffer:

NaPi/NaOH (pH 7,5)	40 mM
--------------------	-------

NaCl	500 mM
------	--------

Elutionspuffer:

NaPi/NaOH (pH 7,5)	40 mM
--------------------	-------

NaCl	500 mM
------	--------

Imidazol	300 mM
----------	--------

Puffer für die Streptavidin-Affinitätschromatographie

SA-Puffer als Laufpuffer:

NaCl	150 mM
------	--------

Tris/HCl pH 8,0	100 mM
-----------------	--------

Elutionspuffer:

D-Desthiobiotin	2,5 mM
-----------------	--------

NaCl	150 mM
------	--------

Tris/HCl pH 8,0	100 mM
-----------------	--------

Regenerationspuffer:

EDTA pH 8,0	1 mM
-------------	------

HABA	5 mM
------	------

NaCl	150 mM
------	--------

Tris/HCl pH 8,0	100 mM
-----------------	--------

SA-Puffer für die Regeneration und Lagerung:

EDTA pH 8,0	1 mM
NaCl	150 mM
Tris/HCl pH 8,0	100 mM
NaN ₃	0,02 % (v/v)

Puffer für die Größenausschlusschromatographie

SEC-Puffer:

NaCl	100 mM
Tris/HCl pH 8,0	50 mM

Puffer für die cytofluorimetrische Analyse

PBS:

KH ₂ PO ₄	4 mM
NaHPO ₄	15 mM
NaC	115 mM

Der pH dieses Puffers liegt bei 7,4.

Sonstige Puffer und Lösungen

Benzonase-Puffer

MgCl ₂	5 mM
Tris/HCl pH 8,0	100 mM

Lagerungslösung für Chromatographie-Säulen:

EDTA pH 8,0	1 mM
NaN ₃	0,02 % (v/v)

Gelöst in ddH₂O und anschließend autoklaviert.

Periplasma-Extraktionspuffer:

EDTA pH 8,0	1 mM
Saccharose	500 mM
Tris/HCl pH 8,0	100 mM

TE-Puffer:

EDTA pH 8,0	1 mM
Tris/HCl pH 8,0	10 mM

Der 1/10 TE-Puffer wurde durch 10-fache Verdünnung des TE-Puffers mit ddH₂O hergestellt.

2.2 Molekularbiologische Methoden

Etablierte Standardarbeitsvorschriften des Lehrstuhls für Biologische Chemie der TU München wurden an frühere Dissertationen der Arbeitsgruppe angelehnt, aktualisiert und ergänzt (Schlapschy, 2002; Friedrich, 2017; Schlehuber, 2001; Deuschle, 2020)

2.2.1 Kultivierung und Konservierung von *E. coli*-Stämmen

Für den Erhalt von *E. coli* Einzelkolonien wurden diese von den LB-Agarkulturplatten gepickt, auf einer neuen LB-Agarkulturplatte ausgestrichen und bei 37 °C über Nacht inkubiert. Durch Zugabe der entsprechenden Antibiotika zum Nährmedium wurde die Selektion auf *E. coli* Zellen, die Plasmide tragen, gewährleistet. Die Kulturplatten mit den Bakterienkulturen wurden bei 4 °C gelagert und maximal 4 Wochen zur Inokulation von Flüssigmedien verwendet.

Flüssigkulturen, angesetzt mit hitzesterilisiertem LB-Medium und dem entsprechenden Antibiotikum, wurden im Volumen von 5 ml im 13 ml Kulturröhrchen oder im 50 ml-Erlenmeyerkolben sowie von 50 ml, 200 ml oder 2 L im 100 ml-, 1 L- bzw. 5 L-Erlenmeyerkolben mit Verschluss aus Aluminiumfolie verwendet. Vorkulturen mit einem Volumen von 5 ml oder 50 ml wurden mit einer Einzelkolonie, stammend von LB-Agarkulturplatten, angeimpft. Die 2 L Flüssigkulturen wurden im Verhältnis 1:100 aus einer stationären 50 ml Vorkultur inokuliert. Sofern nicht anders angegeben, erfolgte die Kultivierung der 5 ml Flüssigkulturen bei 37 °C und die Kultivierung der 50 ml Flüssigkulturen bei 30 °C, jeweils bei 180 rpm. Die 2 L Kultur wurde bei 22 °C und 180 rpm im Schüttelinkubator inkubiert. Die Überprüfung der Zelldichte von Flüssigkulturen erfolgte durch Messung der optischen Dichte bei 550 nm (OD_{550}) in einer Halb-Mikro Küvette aus Kunststoff mit Hilfe eines Spektralphotometers. Die Messwerte wurden gegen das verwendete Medium abgeglichen.

Um eine mit Plasmiden transformierte *E. coli* Flüssigkultur dauerhaft zu konservieren wurde 0,8 ml der Kultur mit 0,25 ml Hitze-sterilisiertem 75 % (v/v) Glycerin gemischt und bei -80 °C gelagert.

2.2.2 Transformation von *E. coli* mit Plasmid-DNA nach der $CaCl_2$ -Methode

Für die Aufnahme von Plasmid DNA müssen *E. coli*-Zellen zuvor in eine transformationskompetente Form überführt werden. Dazu wurde die klassische $CaCl_2$ -Methode (Cohen *et al.*, 1972) eingesetzt, wodurch je nach *E. coli* Stamm ca. 10^5 bis 10^6 Transformanden pro μg Plasmid-DNA erhalten wurden.

Zur Herstellung der kompetenten *E. coli* Zellen wurde eine 50 ml LB-Flüssigkultur im Verhältnis 1:100 mit einer stationären Übernachtskultur des zu transformierenden *E. coli* Stamms angeimpft und bei 37 °C inkubiert. Nach Erreichen einer optischen Dichte von

OD₅₅₀ = 0,5 wurde die Kultur in ein steriles 50 ml Röhrchen überführt und die Zellen durch Zentrifugation (4416 g, 4 °C, 10 min) vom Medium getrennt. Das Medium wurde dekantiert und das Zellsediment in 40 ml eiskalter 100 mM MgCl₂-Lösung resuspendiert. Die Zellen wurden durch erneute Zentrifugation (4416 g, 4 °C, 10 min) sedimentiert und im Anschluss von der MgCl₂-Lösung getrennt. Das Sediment wurde in 20 ml eiskalter 50 mM CaCl₂-Lösung aufgenommen und für 30 min auf Eis inkubiert. Im Anschluss wurden die Zellen ein letztes Mal zentrifugiert (4416 g, 4 °C, 10 min), die CaCl₂-Lösung entfernt und das Sediment in 2 ml 50 mM CaCl₂, 15 % (v/v) Glycerin aufgenommen. Je 100 µl der Zellsuspension mit den kompetenten *E. coli* Zellen wurden in sterilen 1,5 ml Reaktionsgefäße aliquotiert und bei -80 °C bis zur Transformation eingefroren. Alle eingesetzten Lösungen wurden im Autoklaven sterilisiert und vor der Verwendung auf 4 °C abgekühlt.

Für die Transformation wurde ein Aliquot mit 100 µl auf Eis aufgetaut und mit 100 bis 200 ng gelöster Plasmid-DNA oder 7 µl des Ligierungsansatzes (Abschnitt 2.3.4) gemischt und für 30 min auf Eis inkubiert. Anschließend wurde der Ansatz für 45 s bei 42 °C im Heizblock erhitzt, für eine Minute auf Eis gekühlt und mit 900 µl LB-Medium versetzt. Die Zellsuspension wurde zur Regeneration der Zellen für 45 min bei 37 °C unter Schütteln bei 180 rpm inkubiert. Im Anschluss wurde 100 µl davon auf LB-Agar, versetzt mit dem entsprechenden Antibiotikum, ausplattiert und über Nacht bei 37 °C im Brutschrank inkubiert.

2.2.3 Transformation mittels Elektroporation

Zur Generierung der Genbibliotheken wurde eine Transformationseffizienz von mindestens 10⁸/µg DNA benötigt. Dafür wurden *E. coli* BL 21 oder *E. coli* MG1655 Zellen nach dem Protokoll von (Tung & Chow, 1995), in elektrokompetente Form überführt.

Dazu wurde eine 50 ml LB-Flüssigkultur mit einer frisch ausgestrichenen Einzelkolonie des *E. coli* Stammes inokuliert und über Nacht bei 30 °C, 180 rpm im Schüttelinkubator inkubiert. Mit dieser Vorkultur wurde eine 2 L TB-Flüssigkultur im Verhältnis 1:100 angeimpft und bei 30 °C, 105 rpm bis zu OD₅₅₀ = 0,3 - 0,4 inkubiert. Der Kulturkolben wurde anschließend für 20 min unter gelegentlichem Schütteln auf Eis abgekühlt. Die abgekühlte Kultur wurde auf vier sterile SLA-3000 Becher verteilt und zentrifugiert (2700 g, 10 min, 4 °C). Der Überstand wurde vollständig abgenommen und das Zellsediment in jedem der beiden Becher mit 500 ml ddH₂O resuspendiert. Darauf folgte eine erneute Zentrifugation für 13 min bei 2700 g und 4 °C. Der Überstand wurde vollständig abgenommen und das Zellsediment ein zweites Mal mit je 500 ml ddH₂O gewaschen und für 15 min bei 3500 g und 4 °C zentrifugiert. Im Anschluss wurden die Zellen ein letztes Mal mit ddH₂O und 10 % Glycerin gewaschen und bei 3500 g, 4 °C für 15 min zentrifugiert. Nach dieser Zentrifugation wurde der Überstand sorgfältig abgezogen und das Zellsediment in 400 µl GYT-Medium resuspendiert und bis zum Einsatz für die anschließende Transformation auf Eis gelagert.

Für die Transformation von komplexen Gen-Bibliotheken wurde je 100 µl der Zellsuspension mit 1 µg ligierter Plasmid-DNA (Abschnitt 2.3.4) in einer Elektroporationsküvette gemischt und mit Hilfe des MicroPulser Elektroporators (BioRad) im Programm Ec2 (5 ms, 12.5 kV/cm) transfiziert. Direkt nach der Transformation wurde die Zellsuspension in zwei Schritten mit 1 ml SOC Medium resuspendiert und zur Regeneration der Zellen für 45 min bei 37 °C unter Schütteln bei 180 rpm inkubiert. Im Anschluss wurden 100 µl davon für die Bestimmung der Anzahl an Transformanten sowie die restlichen 1,9 ml der Zellsuspension auf LB-Agar, versetzt mit dem entsprechenden Antibiotikum, ausplattiert und über Nacht bei 37 °C im Brutschrank inkubiert.

2.2.4 DNA-Isolierung aus *E. coli*

2.2.4.1 Isolierung von chromosomaler DNA

Für die Isolierung von chromosomaler DNA aus *E. coli* MG1655 wurde 1,5 ml einer stationären Übernachtskultur bei Raumtemperatur und 14.000 rpm sedimentiert (Eppendorf-Zentrifuge 5415C). Das Medium wurde abgezogen und die Zellen in 567 µl TE-Puffer resuspendiert. Nach Zugabe von 30 µl 10 % (w/v) SDS, 2 µl DNase-freier RNase (10 mg/ml) und 3 µl Proteinase K (20 µg/ml) wurde die Suspension mehrfach invertiert und für 1 h bei 37 °C inkubiert. Im Anschluss wurde die Suspension schrittweise mit 100 µl 5 M NaCl und 80 µl 0,7 M NaCl, 10 % (w/v) CTAB versetzt, wobei nach jeder Zugabe die Suspension gemischt wurde. Nach 10-minütiger Inkubation bei 65 °C erfolgte die Abtrennung der Proteine. Dafür wurde die Suspension zuerst mit ml Phenol/Chloroform und danach mit 1 ml Chloroform/Isoamylalkohol versetzt und kräftig gemischt. Die wässrige Phase wurde in ein frisches Reaktionsgefäß überführt und zur Fällung der chromosomalen DNA das 0,6-fache Volumen Isopropanol (p. a.) zugegeben. Nach 90-minütiger Inkubation bei -20 °C wurde die chromosomale DNA mittels Zentrifugation für 10 Minuten bei Raumtemperatur und 14.000 rpm präzipitiert. Das Präzipitat wurde mit 500 µl eiskaltem 70 % (v/v) Ethanol gewaschen und in der Speed-Vac (RVC 2-18 CDplus, Christ) getrocknet. Die DNA wurde in 100 µl TE- Puffer gelöst und bei -20 °C gelagert.

2.2.4.2 Isolierung von doppelsträngiger Plasmid-DNA

Für die Isolierung von Plasmid-DNA aus *E. coli* wird ein Verfahren verwendet, das auf der alkalischen Lyse der Bakterienzellen (Birnboim & Doly, 1979) basiert, gefolgt von der selektiven Immobilisierung der Plasmid-DNA an einem Trägermaterial (z.B. Anionenaustauscher, Silicagel), mehreren Waschschritten und der abschließenden Elution der DNA. Dafür wurde ein entsprechender Kit der Firma Qiagen nach Angaben des Herstellers verwendet.

Für die Analyse von neu konstruierten Plasmiden oder Plasmiden, die durch ortsgerichtete Mutagenese verändert wurden, wurde eine 5 ml-Flüssigkultur in sterilen Röhrchen mit einer Einzelkolonie der Transformanden angeimpft und die Plasmid DNA nach Herstellerangaben präpariert. Bei der Präparation wurden üblicherweise 10 µg DNA erhalten, die für die Restriktions- oder Sequenzanalyse verwendet wurde. Die Plasmid-DNA wurde bei -20 °C gelagert.

Um größere DNA-Mengen für den präparativen Restriktionsverdau zu gewinnen, wurde die Plasmid-DNA aus 5 ml Kultur, stammend aus einer stationären 50 ml Übernachtskultur, isoliert. Die DNA wurde mit Hilfe des Qiagen Plasmid Mini-Präp Kits gemäß Herstellerangaben präpariert und die Ausbeute betrug in der Regel ca. 25 µg DNA. Bis zur Verwendung wurde die DNA bei -20 °C gelagert.

2.2.5 Gelelektrophorese und Reinigung von DNA

Für präparative wie auch für analytische Zwecke wurde eine größenabhängige Trennung von DNA-Fragmenten mittels Agarose-Gelelektrophorese in Gegenwart von Ethidiumbromid (Sambrook *et al.*, 1989) durchgeführt.

Abhängig von der Größe der zu analysierenden DNA-Fragmente wurden Agarosegele mit Konzentrationen von 0,5 – 2 % Agarose in TAE-Puffer hergestellt. Dafür wurde eine Suspension von Agarose in TAE-Puffer, bis die Agarose vollständig gelöst war, erhitzt und mit 1/10.000 Volumen einer 1 % (w/v) Ethidiumbromidlösung versetzt. Diese Lösung wurde in eine horizontale Gelkammer mit Taschenformer gegossen und bei Raumtemperatur bis zum Erstarren des Gels abgekühlt. Für Gelelektrophoresen mit dem Ziel der präparativen Isolierung von DNA-Fragmenten wurde *Low-Melting-Point*-Agarose („UltraPure“, Invitrogen) mit einem erhöhten Reinheitsgrad verwendet. Das Gel wurde in eine mit TAE-Puffer gefüllte Elektrophoresekammer gelegt. Die DNA-Proben wurden mit einem 6x-DNA-Auftragspuffer versetzt und in die Geltaschen überführt. Die Elektrophorese wurde bei einer konstanten Spannung von 90 V, im Fall von niedrig-schmelzenden präparativen Gelen bei 70 V, durchgeführt. Anschließend wurden die DNA-Fragmente anhand der Fluoreszenz des in die DNA interkalierenden Ethidiumbromids bei einer Bestrahlung mit UV-Licht ($\lambda = 312 \text{ nm}$) sichtbar gemacht und unter Verwendung eines Rotfilters fotografiert. Die Größe der Fragmente wurde mit Hilfe eines DNA-Standards abgeschätzt und zur Präparation der gewünschten DNA-Fragmenten wurde die entsprechende Bande mit einem Skalpell ausgeschnitten. Zur Reinigung der isolierten DNA von der Agarose für anschließende Klonierungsarbeiten wurde der Qiagen Gel Reinigungs Kit nach der Gebrauchsanweisung des Herstellers verwendet. Die Bestimmung der DNA-Konzentration erfolgte mit Hilfe des NanoDrop Photometers anhand der UV-Lichtabsorption bei 260 nm unter Verwendung des Lambert-Beerschen Gesetzes. Der

entsprechende Extinktions-Koeffizient wurde ausgehend von einem Durchschnittswert von $660 \text{ M}^{-1} \text{ cm}^{-1}$ pro Basenpaar berechnet.

2.2.6 Gen-Deletion mit Hilfe der λ -Red Rekombination

Für die Deletion eines Gens wurden *E. coli* NEB Express Zellen, die das Plasmid pKD46 trugen (Datsenko & Wanner, 2000), in eine elektrokompente Form überführt. Dabei wurde, abweichend zum Abschnitt 2.2.3, die 2 L-Hauptkultur mit 10 mM L-Arabinose versetzt, um die Expression der auf dem Plasmid pKD46 kodierten Lambda-Red Proteine (Gam, Beta, Exo) zu induzieren. Ein Aliquot, diesmal bestehend aus 50 μl elektrokompenter *E. coli* Zellen, wurden mit 400 ng Donor-DNA (Kanamycin Sequenz flankiert von 500 bp langen und zur Ziel-DNA homologen Armen) in einer MicroPulser Küvette (5 ms, 12.5 kV/cm) transfiziert. Nach der Regenerierung der transfizierten *E. coli* NEB Express Zellen für 1 h bei 37 °C wurden diese auf Selektionsplatten mit 35 mg/L Kanamycin ausplattiert, welche bei 37 °C über Nacht inkubiert wurden. Durch die Inkubation bei 37 °C wurden die *E. coli* Zellen von dem Plasmid pKD46, das einen temperatursensitiven Replikationsursprung (pSCS101) trägt, kuriert. Solchen Klonen war es nicht mehr möglich, auf Agarplatten mit 100 mg/L Ampicillin zu wachsen, da das resistenzvermittelnde Gen ausschließlich auf diesem Plasmid kodiert war. Die Deletion wurde mittels Colony-PCR bestätigt.

2.2.7 Gen-Modifizierung mit Hilfe der Cas9 / λ -Red Rekombination

Die Gen-Modifizierung erfolgte mit dem Plasmidsystem pCas (Addgene 62225) und pTargetF (Addgene 62226), welches sowohl die homologe Rekombination mit einer Donor-DNA als auch die anschließende Selektion ermöglichte (Jiang *et al.*, 2015). Auf pCas sind die Endonuklease Cas9 für die Selektion sowie die Lambda-Red Proteine Gam, Beta und Exo für die Rekombination mit der Donor-DNA kodiert. pTargetF kodiert für die sgRNA Gensequenz, deren N20-Spacer spezifisch für den chromosomalen Ziellokus der Rekombination gewählt wurde und in Kombination mit Cas9 bei der unveränderten Originalsequenz einen Doppelstrangbruch auslöst.

Die zu modifizierenden *E. coli* MG1655 Zellen wurden mit dem Plasmid pCas mittels Elektroporation transfiziert, welche daraufhin die unter einem konstitutiven Promoter stehende Endonuklease Cas9 exprimierten. Ausgehend von einer transformierten Kolonie wurden die *E. coli* Zellen in eine elektrokompente Form überführt. Dabei wurden in Abweichung von Abschnitt 2.2.3 die auf pCas kodierten Lambda-Red Proteine durch Zugabe von 10 mM L-Arabinose in einer 2 L-Hauptkultur induziert, welche bei OD_{600} von 0,7 geerntet wurde. 50 μl der hergestellten elektrokompenten Zellsuspension wurde mit 100 ng pTargetF und 400 ng Donor-DNA (3.3.2.1) in einer MicroPulser Küvette transfiziert (5 ms, 12.5 kV/cm). Die transfizierten Zellen wurden in 1 ml SOC-Medium bei 30 °C für 1 h regeneriert und auf LB-

Agarplatten mit Kanamycin für die Selektion auf pTarget und Spectinomycin für die Selektion auf pCas ausplattiert. Nach 43-stündiger Inkubation bei 30 °C wurden Kolonien gepickt und mittels analytischer *Colony-PCR* hinsichtlich erfolgreicher Rekombination untersucht. Die finale Überprüfung erfolgte durch Sequenzierung des *Colony-PCR*-Produktes.

Entsprechende *E. coli* Kolonien wurden daraufhin schrittweise von den Plasmiden pCas und pTargetF kuriert. Dafür wurde eine Einzelklon gepickt und in einer 2-ml LB-Flüssigkultur bei 30 °C kultiviert. Die Induktion der Translation der auf pCas kodierten und zum Replikationsursprungs des Plasmides pTargetF homologe sgRNA erfolgte direkt bei der Inokulation mit 500 µM IPTG. Der N20 Spacer hybridisierte dabei im Replikationsursprung von pTargetF und induzierte einen Doppelstrangbruch durch Cas9. Zur Überprüfung wurde 100 µl und 1 ml der Zellsuspension auf LB-Agarplatten mit Spectinomycin ausgestrichen und gegeben falls vereinzelt. 10 -20 Einzelklone wurden im Anschluss jeweils auf einer Agarplatte mit Spectinomycin und einer Agarplatte mit Kanamycin transferiert. Von pTargetF erfolgreich kurierte *E. coli* Zellen konnten auf Agarplatten mit Kanamycin nicht wachsen. Um von pTargetF kurierte Kolonien im zweiten Schritt von pCas zu kurieren, wurden Kolonien gepickt und bei 37 °C über Nacht in LB-Medium ohne Antibiotikum inkubiert. Am darauffolgenden Tag wurde aus der Flüssigkultur *E. coli* durch einen Vereinzelausstrich auf Agarplatten, denen kein Antibiotikum zugegeben wurde, transferiert und über Nacht bei 37 °C inkubiert. Das wurde so lange wiederholt, bis die Zellen vom Plasmid pCas, welches einen aufgrund eines temperatursensitiven Ursprungs, der über 30 °C nicht repliziert wird, kuriert waren. Dies wurde überprüft indem eine Kolonie von einer Agarplatte ohne Antibiotikum auf Agarplatten mit Spectinomycin oder Kanamycin transferiert wurde und auf diesen kein Wachstum zeigte.

2.3 *In-vitro*-Modifizierung von DNA

2.3.1 Spaltung doppelsträngiger DNA mit Restriktionsendonukleasen

Plasmid-DNA wurde zur Präparation von DNA-Fragmenten, sowie zur Restriktionsanalyse mit den entsprechenden Restriktionsendonukleasen gespalten. Dafür wurden die vom Hersteller empfohlenen Angaben bezüglich Puffer, Temperatur und Reaktionszeit eingehalten. Sollten zwei oder mehr Restriktionsendonukleasen verwendet werden, wurde derjenige Puffer gewählt, in dem alle Enzyme möglichst hohe Aktivität besaßen. Ein präparativer Verdau wurde typischerweise mit 2 µg DNA und 20 U jedes benötigten Restriktionsenzyme durchgeführt, außer für Bibliotheken, wofür 20 µg DNA und 200 U der Restriktionsenzyme eingesetzt wurden. Da nur Restriktionsenzyme ohne Sternaktivität verwendet wurden, wurde die Reaktionszeit zwischen 3 h und bis zu 16 h gewählt. Bei einem analytischen Verdau wurde 500 ng Plasmid DNA mit 5 U der Restriktionsendonuklease in einem 10 µl Reaktionsansatz für 1 h inkubiert. Die Inaktivierung erfolgte enzymspezifisch je nach den Angaben des Herstellers.

2.3.2 Phosphorylierung und Dephosphorylierung von DNA

Um nach einem präparativen Verdau eine intramolekulare Rezirkularisierung bei der Ligierung von DNA-Fragmenten zu verhindern, wurde das Vektorfragment mit Alkalischer Phosphatase (Fast AP) dephosphoryliert. Dazu wurde nach dem Restriktionsverdau des Vektors 1 U Fast AP je μg verdauter DNA zugeben und 1 h bei 37 °C inkubiert. Das Enzym wurde durch 15-minütige Inkubation bei 65 °C inaktiviert und das dephosphorylierte DNA-Fragment mittels präparativer Agarose-Gelelektrophorese (Abschnitt 2.2.5) isoliert.

DNA-Fragmente, die mittels Hybridisierung von Oligodesoxynukleotiden erzeugt wurden, wurden an den 5'-Enden mit Hilfe einer T4-Polynukleotid-Kinase (PNK) phosphoryliert. Hierzu wurde das Fragment laut den Herstellerangaben in PNK-Puffer mit 10 mM ATP und 10 U PNK pro 300 pmol 5' DNA Enden für 30 min bei 37 °C inkubiert. Das Enzym wurde für 20 min bei 65 °C zur Inaktivierung inkubiert.

2.3.3 Polymerase-Kettenreaktion (PCR) zur Amplifizierung von Genen

Die Polymerase-Kettenreaktion (*Polymerase Chain Reaction*, PCR; (Saiki *et al.*, 1988)) wurde in der vorliegenden Arbeit zur selektiven Amplifizierung von Genen, zur ortsgerichteten Mutagenese und zur Einführung von Zufallsmutationen (*Error Prone PCR*) eingesetzt.

2.3.3.1 PCR-Amplifizierung eines Gens

Die zu amplifizierende Gensequenz wurde von zwei Oligodesoxynukleotiden (*Primer*) flankiert, die an ihrem 3'-Ende eine mindestens 18 Basen lange, zur Matrizen-DNA komplementäre Sequenz aufwies. Für die PCR wurde die Hot Start *Taq* DNA-Polymerase oder die Q5 High-Fidelity DNA-Polymerase, die sich aufgrund ihrer *Proofreading*-Aktivität durch eine niedrige Fehlerrate auszeichnet, verwendet.

Für eine Standard PCR oder eine *Colony*-PCR wurde die Hot Stat *Taq* DNA-Polymerase, eine Weiterentwicklung der *Taq* DNA-Polymerase (Chien *et al.*, 1976) verwendet. Typischerweise enthielt ein 50 μl -PCR-Ansatz 1 bis 100 ng Matrizen-DNA, je 0,1 μM der beiden *Primer*, 5 μl 10x Hot Start *Taq* Puffer, 1 μl dNTP-Mix (je 10 mM dATP, dCTP, dGTP, dTTP) und 1,25 U Hot Start *Taq* Start DNA-Polymerase. Nach initialer Denaturierung für 2 min bei 95 C erfolgte die PCR in drei aufeinanderfolgenden Zyklen: 1. Denaturierung der doppelsträngigen DNA bei 95 °C für 1 min; 2. Hybridisierung (*Annealing*) der *Primer* an die DNA-Matrix bei 50 bis 60 °C (je nach Schmelztemperatur des gebildeten Hybrids) für 30 s; 3. Synthese der DNA bei 72 °C für 1 min. Der Zyklus wurde in der Regel 30-mal durchlaufen und der Ansatz schließlich zur Vervollständigung der Syntheseprodukte für 5 min bei 72 °C inkubiert. Im Gegensatz zu der Standard PCR diente bei einer *Colony*-PCR die DNA einer lysierten Einzelkolonie als Matrize. Dafür wurde eine Kolonie von einer Platte gepickt und direkt im Reaktionsmix resuspendiert. Die Zellen lysierten während der initialen Denaturierung. Die *Colony*-PCR wurde für

analytische Zwecke meist im 10 µl-Maßstab durchgeführt der je 0,1 µM der beiden *Primer*, 1 µl 10x Hot Start *Taq* Puffer, 0,2 µl dNTP-Mix (je 10 mM dATP, dCTP, dGTP, dTTP) und 0,25 U Hot Start *Taq* Start DNA-Polymerase beinhaltete.

Für die Amplifizierung einer Gensequenz mit der Q5 High-Fidelity DNA-Polymerase enthielt ein 50 µl-Standardansatz 1 bis 10 ng Matrizen-DNA, 0,5 µM der jeweiligen *Primer*, je 200 µM der vier dNTPs, 10 µl des 5x Q5-Puffers und 1 U Q5 High-Fidelity DNA-Polymerase. Nach initialer Denaturierung für 2 min bei 98 °C erfolgte die PCR wieder in drei Schritten: 1. Denaturierung der doppelsträngigen DNA bei 98 °C für 30 s; 2. Hybridisierung (*Annealing*) der *Primer* an die DNA-Matrix bei 50 bis 65 °C (je nach Schmelztemperatur des gebildeten Hybrids) für 15 s; 3. Synthese der DNA bei 68 °C für 1 min. Der Zyklus wurde in der Regel 30-mal durchlaufen und der Ansatz schließlich zur Vervollständigung der Syntheseprodukte für 2 min bei 68 °C inkubiert.

Um Kontaminationen ausschließen zu können, wurde ein Reaktionsansatz ohne Matrizen-DNA ebenfalls durchgeführt. Ein Aliquot von 5 µl des erhaltenen PCR Produktes wurde für ein analytisches Agarosegel verwendet, das restliche PCR Produkt wurde mit dem QIAquick Gel Extraktions Kit nach den Vorgaben des Herstellers gereinigt.

2.3.3.2 Ortsgerichtete Sättigungsmutagenese und *Overlap-Extension* PCR

Die Randomisierung von Aminosäuren an einer bestimmten Position wurde durch die Verwendung von partiell degenerierten Oligodesoxynukleotiden mit ein oder zwei NNS-Codons an den zu randomisierenden Aminosäurepositionen erreicht. In zwei parallelen PCR-Ansätzen wurden zwei DNA-Abschnitte amplifiziert, die sich in der Nähe der Mutagenesepositionen am 3' bzw. 5' Ende überlappten. Die beiden PCR Produkte wurden durch präparative Gelelektrophorese und mittels QIAquick Gel Extraktions Kit gereinigt. Beide Fragmente wurden in der darauffolgenden *Overlap-Extension* PCR (Ho *et al.*, 1989) als Matrizen DNA eingesetzt, zusammen mit den flankierenden *Primern*. Nach der initialen Denaturierung der Matrizen-DNA hybridisierten die überlappenden Enden und durch die PCR ein entstand ein zusammenhängendes DNA-Fragment. Nach einer präparativen Gelelektrophorese und Reinigung des PCR Produktes diente dieses als Matrize für eine finale PCR mit den flankierenden *Primern*. Die Bedingungen für die hier durchgeführte PCR mit der Q5 High-Fidelity DNA-Polymerase waren analog zu Abschnitt 2.3.3.1.

2.3.3.3 Ortsgerichtete Mutagenese durch QuikChange PCR

Der zielgerichtete Austausch von 1 bis 3 Basen innerhalb eines Bereichs von ca. 50 Basenpaaren erfolgte unter Verwendung des QuikChange SiteDirected Mutagenesis Kits (Zheng *et al.*, 2004).

Dieses Verfahren beruht auf der Synthese vollständiger Plasmide durch eine PCR mit der PfuUltra II Fusion Hot Start (HS) DNA-Polymerase und zwei komplementären *Primern*, welche die entsprechende Substitution, Insertion oder Deletion enthalten. Als Matrize diente Plasmid-DNA, die zuvor aus einem *dam*⁺ *E. coli* Stamm (z.B. XL1-Blue) isoliert wurde. Ein Standardansatz von 50 µl enthielt 10 ng Matrizen-DNA, 0,1 µM der jeweiligen *Primer*, je 200 µM der vier dNTPs, 10 µl des 5x *Pfu* Ultra-Reaktionspuffers und 2,5 U *Pfu* Ultra II Fusion HS DNA-Polymerase. Nach einer initialen Denaturierung für 30 s bei 95 °C erfolgte die PCR wieder in drei Schritten: 1. Denaturierung der doppelsträngigen DNA bei 95 °C für 30 s; 2. Hybridisierung (*Annealing*) der *Primer* an die DNA-Matrix bei 54 °C für 1 min; 3. Synthese der DNA bei 72 °C für 5 min und 30 s. Der Zyklus wurde 18-mal durchlaufen und der Ansatz schließlich zur Vervollständigung der Syntheseprodukte für 2 min bei 72 °C inkubiert. Der Erfolg der PCR wurde durch analytische Agarose-Gelelektrophorese überprüft. Parentale, noch methylierte DNA wurde daraufhin durch Zugabe von 20 U *DpnI* zum PCR-Ansatz bei 37 °C für eine Stunde verdaut. Mit 7 µl des PCR Produktes wurden chemisch kompetente Zellen transfiziert, die Plasmid-DNA aus einem Einzelklon isoliert und mittels Sequenzierung überprüft.

2.3.3.4 Zufallsmutagenese mittels *Error Prone PCR*

Für die Zufallsmutagenese eines definierten Genabschnitts zur Herstellung von Gen-Bibliotheken wurde das GeneMorph II-Mutagenesekit von Agilent verwendet. Dieses Kit besteht unter anderem aus einem Gemisch aus zwei rekombinanten DNA-Polymerasen, die den zum Matrizen-Strang komplementären DNA-Strang mit erhöhter Fehlerrate synthetisiert (Hogrefe & Cline, 2001; Hogrefe & Hansen, 1997; Hogrefe, 1997). Um die gewünschte Mutationsrate von 0 bis 4 Nukleotidaustausche für ein DNA-Fragment von 606 bp zu erreichen, wurde 200 ng der zu amplifizierenden Sequenz als Matrizen-DNA eingesetzt. Die initiale Denaturierung der doppelsträngigen DNA erfolgte für 2 min bei 95 °C, gefolgt von drei 12-mal wiederholten Schritten: 1. Denaturierung der doppelsträngigen DNA bei 95 °C für 30 s; 2. Hybridisierung (*Annealing*) der *Primer* an die DNA-Matrix bei 59 °C für 30 s; 3. Synthese der DNA bei 72 °C für 1 min. Der Ansatz wurde schließlich zur Vervollständigung der Syntheseprodukte für 10 min bei 72 °C inkubiert. Der Erfolg der PCR wurde durch analytische Agarose-Gelelektrophorese überprüft. Das PCR-Produkt wurde mit dem QIAquick Gel Extraktions Kit gereinigt, mit geeigneten Restriktionsendonukleasen verdaut und schließlich erneut mit dem QIAquick Gel Extraktions Kit gereinigt.

2.3.3.5 Amplifizierung der *MjTyrRS* Zufalls-Bibliothek

Die SlonoMax Genbibliothek (dankend erhalten von Sloning BioTechnology, Puchheim, jetzt Teil von MorphoSys, Planegg) basiert auf der naiven *MjTyrRS* bei der 17 Aminosäurepositionen teilrandomisiert wurden. So befand sich nach der Synthese an Position

32, 65, 155 und 158 jede Aminosäure mit 5 %iger Wahrscheinlichkeit. An Position 67, 108, 109, 159, 162 und 177 waren 50 % der Aminosäurereste konserviert und die verbliebenen 50 % randomisiert, an Position 103, 114, 154, 166 und 168 waren 90 % konserviert und die verbliebenen 10 % randomisiert, was bei gleichmäßiger Verteilung der restlichen 19 Aminosäuren zu einem Randomisierungsgrad von 2,5 % bzw. 0,5 % führte. Daraus ergab sich eine Komplexität von 10^{22} . Für die Klonierung dieser Bibliothek auf dem Vektor pSB4.6c2 oder pSB16.6c2 wurde die erhaltene DNA zunächst amplifiziert. Dafür wurden insgesamt 10 PCRs mit jeweils 6 ng der DNA-Bibliothek, das in etwa 10^{11} Moleküle der Aminoacyl-tRNA Synthetase-Varianten entsprach, durchgeführt. Die PCR wurde mit der Q5 High-Fidelity DNA-Polymerase analog zu Abschnitt 2.3.3.1 durchgeführt. Die Zyklenzahl wurde hierbei von 30 auf 15 reduziert, um einem Diversitätsverlust durch die übermäßige Amplifikation einzelner DNA-Moleküle entgegenzuwirken. Im Anschluss wurde das gereinigte PCR-Produkt mit *EamI* verdaut und nach erneuter Reinigung mit dem ebenfalls mit *EamI* verdaute Vektorrückgrat (Abschnitt 2.3.1) ligiert (Abschnitt 2.3.4).

2.3.3.6 Klonierung einer Gen-Bibliothek mittels Sexual PCR (Gene Shuffling)

Für die Generierung einer Gen-Bibliothek mittels Sexual-PCR, auch DNA-*Shuffling* genannt (Stemmer, 1994), wurden die Gene für die drei künstlichen Aminoacyl-tRNA Synthetasen, AzAlaRS (Baumann *et al.*, 2019b), PhotoDopaRS (Hauf *et al.*, 2017) und HqAlaRS (Liu *et al.*, 2013) als doppelsträngige DNA-Fragmente von der Firma Eurofins (Ebersberg, Deutschland) synthetisiert. Dabei wurden die verwendeten Codons weitgehend in Übereinstimmung mit dem synthetischen Gen für die BpaRS (Edwardraja *et al.*, 2017) gewählt, damit die Homologie der Gene größer als 85 % war, was für die Methode essentiell ist. Die Gene wurden zunächst mittels PCR unter Verwendung der *Primer* ep1 und ep2 und der Hot Start *Taq* Hot Start DNA-Polymerase amplifiziert (Abschnitt 2.3.3.1).

Nach der Reinigung wurden jeweils 3,4 µg des PCR-Produktes vereint und die insgesamt 10 µg DNA in einem 50 µl-Ansatz mit 0,7 U DNaseI in einem Puffer aus 2,5 mM $MnCl_2$, 0,5 mM $CaCl_2$ und 10 mM Tris/Cl, pH 7,6 zu 30 – 200 bp langen Fragmenten verdaut. Der Verdau erfolgte sowohl für 3 sowie für 4 min bei 37 °C und wurde durch Zugabe von 12,4 µl eiskalter Stop-Lösung (50 mM EDTA, 30 % (v/v) Glycerin und einer Spatelspitze an Bromphenolblau) inhibiert (Volkov & Arnold, 2000), gefolgt von der thermischen Denaturierung der DNaseI bei 95 °C für 10 min. Die Fragmente wurden ihrer Länge nach mittels präparativer Agarose-Gelelektrophorese (2 % Agarosegel) getrennt, wobei jene Fragmente, die im Größenbereich zwischen 30 und 200 bp lagen, isoliert und gereinigt wurden. Chimäre DNA-Fragmente wurden durch eine PCR-ähnliche Reaktion ohne *Primer* generiert, die mit 20 ng/µl der verdauten und gereinigten DNA-Fragmente, jeweils 10 mM dNTPs und 1,25 U Hot Start *Taq* DNA-Polymerase für 45 Zyklen durchgeführt wurde. Als *Annealing*-Temperatur wurde 50 °C gewählt. Nach einer erneuten Reinigung der Bande bei ca. 600 bp erfolgte eine

finale PCR mit den *Primern* ep1 und ep2 (Abschnitt 2.3.3.1). Im Anschluss erfolgte eine *Error Prone* PCR mit dem gereinigten Produkt der finalen PCR als Matrize und den *Primern* ep1 und ep2 zur Diversifizierung der Gen-Bibliothek. Das Produkt der *Error Prone* PCR wurde mit *EamI* verdaut (Abschnitt 2.3.1) und mit dem ebenfalls verdauten und zusätzlich dephosphorylierten Vektor pSB16.6c2 ligiert (Abschnitt 2.3.4).

2.3.4 Ligierung von DNA-Fragmenten

Die kovalente Verknüpfung von mit Restriktionsendonukleasen verdauten DNA-Fragmenten und Vektoren zu vollständigen Plasmiden erfolgte mit Hilfe der T4 DNA-Ligase.

Für eine Standard-Ligierung mit einem Gesamtvolumen von 20 µl wurden 2 µl 10x Puffer (250 mM Tris-HCl pH 7,6; 50 mM MgCl₂; 5 mM ATP; 5 mM Dithiothreitol; 25 % (w/v) Polyethylenglycol 8000, 1 µl T4 DNA-Ligase (1 U/µl) und insgesamt 200 ng DNA mit einem Vektor zu Insert Verhältnis von 1 zu 3 eingesetzt. Der Ansatz wurde entweder für 1 h bei Raumtemperatur oder über Nacht bei 16 °C im Wasserbad inkubiert. Zur Kontrolle wurde 200 ng verdaute Plasmid-DNA ohne Insert in einem separaten Ansatz ebenfalls einer Ligierungsreaktion unterworfen. Im Anschluss an die Inkubation wurden mit jeweils 5 µl des Ligierungsansatzes CaCl₂-kompetente *E. coli*- Zellen transformiert.

Für die Ligierung der Aminoacyl-tRNA Synthetase-Bibliotheken (Abschnitt 2.3.3.5 und 2.3.3.6) wurde das daraus resultierende PCR-Produkt nach dem Verdau mit *EamI* mit dem ebenfalls mit *EamI* verdauten Vektor pSB4.6c2 oder pSB16.6c2 eingesetzt. Dabei wurde die benötigte Menge an DNA anhand der vorliegenden Menge an Aminoacyl-tRNA Synthetase-Fragmenten berechnet, so dass das Vektor zu Insert Verhältnis 1:3 betrug. Das Gesamtvolumen des Ligierungsansatzes ist abhängig von der eingesetzten DNA-Menge: 1 µl pro 10 ng DNA. Die Ligase-Konzentration wurde nach den Angaben des Herstellers mit 0,1 U/µl gewählt. Die Inkubation erfolgte für 72 h bei 16 °C. Die DNA wurde im Anschluss an die Ligierung aus der wässrigen Lösung durch die Zugabe des gleichen Volumens an 5 M Ammoniumacetatlösung, 1/40 des Volumens einer tRNA-Stammlösung (10 mg/ml in TE-Puffer) und das vierfache Volumen an Ethanol (p. a.) bei -20 °C für 2 h präzipitiert. Der Fällungsansatz wurde zentrifugiert (16.000 g, 4 °C, 30 min), der Überstand verworfen und das Präzipitat mit 70 % (v/v) Ethanol gewaschen. Nach erneuter Zentrifugation (16.000 g, 4 °C, 30 min) wurde der Überstand restlos entfernt und das Sediment in der Speed-Vac für 20 min bei 50 °C getrocknet, bevor die gefällte DNA in 5 µl pro 200 ng eingesetzter DNA gelöst wurde. Mit den so generierten Gen-Bibliotheken wurden elektrokompente *E. coli* Zellen transformiert (Abschnitt 2.2.3).

2.3.5 Sequenzierung doppelsträngiger DNA

DNA-Abschnitte, die durch Mutagenese oder eine PCR amplifiziert wurden, wurden mittels Sequenzanalyse überprüft. Dafür wurde Plasmid-DNA oder das gereinigte PCR Produkt vom Anbieter Eurofins Genomics mit Hilfe der der Cycle-Sequenzierung Methode (Kretz *et al.*, 1994), einer Erweiterung der Kettenabbruch-Methode (Sanger *et al.*, 1977; Kretz *et al.*, 1994), verwendet.

2.4 Gentechnische Methoden

2.4.1 Produktion rekombinanter Proteine in *E. coli*

Die gentechnische Herstellung von Borocalin-Varianten erfolgte in *E. coli* BL21. Diese wurden aufgrund der Fusion mit der bakteriellen Signalsequenz (OmpA) in den periplasmatischen Raum von *E. coli* sekretiert (Skerra & Plückthun, 1988). Die Gewinnung des periplasmatischen Extraktes erfolgte durch selektive Permeabilisierung der äußeren Membran nach der Ernte der Zellen. Für die Selektion der Aminoacyl-tRNA Synthetase-Bibliotheken sowie die Evaluierung der *E. coli* Stammvarianten zur Steigerung der Proteinausbeute wurde eGFP(39a) in *E. coli* MG1655 und dessen Derivaten cytoplasmatisch exprimiert.

2.4.1.1 Periplasmatische Borocalin-Produktion in *E. coli* im 2 l-Maßstab

Für die Produktion der Borocalin-Varianten wurden *E. coli* BL21 Zellen mit dem Expressionsplasmid pSB4.6c14, auf dem die gewünschte Borocalin-Variante kodiert war, transformiert. Von den Transformanten wurde eine Einzelkolonie gepickt, eine Vorkultur (50 ml LB-Flüssigkultur, 30 mg/l Chloramphenicol) angeimpft und diese über Nacht bei 30 °C und 180 rpm inkubiert. Die Expression des rekombinanten Proteins erfolgte am nächsten Tag nach Inokulation einer 2 L-Flüssigkultur (LB-Medium, 30 mg/l Chloramphenicol) in einem 5 L-Erlenmeyerkolben, wobei die initiale OD₅₅₀ 0,1 betrug. Die so erzeugte Hauptkultur wurde bei 22 °C und 180 rpm inkubiert. Bei Erreichen der OD₅₅₀ von 0,3 wurde die Kultur mit 1 mM der nicht-natürlichen Aminosäure D/L-Bpa versetzt. Die Genexpression der Aminoacyl-tRNA Synthetase, BpaRS, erfolgte eine Stunde nach der Zugabe von Bpa durch die Induktion mit aTc (50 ng/L). Nach einer weiteren Stunde wurde mit der Zugabe von IPTG (500 µM) die Expression der Borocalin-Variante für 16 h induziert. Die Bakterien wurden geerntet und das durch Zentrifugation erhaltene Zellsediment wurde auf Eis zügig in 20 ml (pro 2 l Kultur) Periplasma-Extraktionspuffer resuspendiert. Nach Inkubation für 30 min bei 4 °C wurden die Sphäroplasten mittels Zentrifugation (5100 rpm, 4 °C, 1 h, Sigma 4-16K) bis zur Klärung des Überstandes abgetrennt, gefolgt von einem weiteren Zentrifugationsschritt (SS34, 15000 rpm, 15 min, 4 °C) zur Abtrennung von Zellbruchstücken. Der daraus resultierende periplasmatische

Extrakt wurde für die Reinigung der Borocalin-Varianten mittels Affinitätschromatographie verwendet (Abschnitt 2.5.1 und 2.5.2).

Zur Analyse der rekombinanten Genexpression durch SDS-PAGE wurde zu unterschiedlichen Zeitpunkten eine Gesamtzellprobe vorbereitet. Dabei wurde 1 ml der Kultur entnommen und die Zellen durch Zentrifugation (13.000 rpm, 4 °C, 5 min) sedimentiert. Die Bakterien wurden, je nach gemessen OD₅₅₀-Wert, mit 100 µl Benzonasepuffer (12,5 U/ml Benzonase) pro OD₅₅₀-Einheit resuspendiert und für 1 h inkubiert. Dabei wurde die freigesetzte Nukleinsäuren durch die Benzonase hydrolysiert. Nach Zugabe von reduzierendem SDS-Auftragspuffers im Verhältnis 1:5 wurde die Probe bei 95 °C für 5 min erhitzt und davon 10 µl auf ein SDS-Gel geladen oder bei -20 °C für eine spätere Analyse eingefroren. Vom periplasmatischen Extrakt wurde ein Probenvolumen von 80 µl entnommen, ebenfalls mit 20 l des reduzierenden SDS-Auftragspuffers gemischt sowie bei 95 °C für 5 min erhitzt und auf das SDS-Gel geladen (Abschnitt 2.5.5).

2.4.1.2 Cytoplasmatische Proteinproduktion in *E. coli* im 5 ml- und 200 ml-Maßstab

Für die Produktion rekombinanter Proteine im Cytoplasma wurden *E. coli* Zellen (MG1655, oder MG1655-Tev2a-c) mit dem entsprechenden Expressionsplasmid transformiert. Eine 5 ml oder 50 ml LB-Flüssigkultur (30 mg/L Chloramphenicol oder 100 mg/L Ampicillin), die als Vorkultur diente, wurde jeweils mit einer Einzelkolonie angeimpft und bei 30 °C und 180 rpm für 16 h inkubiert. Die 50 ml-Hauptkultur (LB-Medium, 30 mg/L Chloramphenicol oder 100 mg/L Ampicillin) für die Expression des Reporterproteins eGFP(39a) zur weiteren Analyse mit dem FACS-Gerät wurde aus der Vorkultur mit einer initialen OD₅₅₀ von 0,1 inokuliert. Nach Erreichen einer OD₅₅₀ von 0,3 unter Verwendung des Expressionsplasmides pSB4.6c2 oder einer OD₅₅₀ von 0,4 unter Verwendung der Expressionsplasmide basierend auf pSB16.6c2, wurden jeweils zweimal 5 ml der Hauptkultur in einen frischen 50 ml Erlenmeyerkolben überführt und die entsprechende, nicht-natürliche Aminosäure oder zur Kontrolle die natürliche Aminosäure mit einer Endkonzentration von 1 mM zugegeben. Nach einer Stunde erfolgte bei den Kulturen mit dem Plasmid pSB4.6c2 die Induktion der Aminoacyl-tRNA Synthetase durch aTc (50 ng/L) und nach einer weiteren Stunde die Induktion des eGFP(39a)-Gens durch 500 µM IPTG. Durch den konstitutiven Promotor proM vor dem Strukturgen für die entsprechende Aminoacyl-tRNA Synthetase, das auf den pSB16.6c2 basierten Plasmiden kodiert war, musste diese nicht separat induziert werden. Daher wurde 1 h nach der Zugabe der jeweiligen nicht-natürlichen Aminosäure die Expression des Reporterproteins eGFP(39a) mittels 200 ng/L aTc induziert. Für die Analyse mit dem FACS-Gerät wurde nach 3-stündiger sowie nach 16-stündiger Expression des Reporterproteins jeweils 1 ml aus dem Schüttelkolben entnommen (Abschnitt 2.6.3).

Die präparative Produktion des Reporterproteins mit den auf pSB16.6c2 basierenden Expressionsplasmiden zwecks Vergleichs der Ausbeuten erfolgte analog in 200 ml LB-Flüssigkultur (100 mg/L Ampicillin). Die nach der 16-stündigen eGFP(39a)-Expression geernteten Zellen wurden in 10 ml SA-Puffer (100 mM Tris/HCl, 500 mM Saccharose, 1 mM EDTA, pH 8,0) resuspendiert und mit Hilfe des Ultraschall-Homogenisators (S250D, 10 s Puls bei einer Amplitude von 30 %, 20 s Pause, Gesamtschallzeit 2 min) bis zur vollständigen Zellyse behandelt. Der erhaltene Zellextrakt wurde durch Zentrifugation (SS-34, 16.000 rpm, 4 °C, 30 min) geklärt und für die Reinigung des Reporterproteins mittels Affinitätschromatographie verwendet (Abschnitt 2.5.1).

Die Analyse hinsichtlich der molekularen Erkennung von Bfa (synthetisiert von Dr. Markus Anneser, Lehrstuhl für Biologische Chemie) durch die BpaRS wurde in M9-Medium (100 mg/L Ampicillin) durchgeführt. Dafür wurde eine 5 ml-Vor-Vorkultur mit einem Einzelkolonie angeimpft und 500 µl davon als Inokulum für die 50 ml-M9-Vorkultur (100 mg/L Ampicillin) verwendet. Die Inokulation der 50 ml-M9-Hauptkultur und die Expression des Reporterproteins erfolgte analog zu der Expression in LB-Medium, abgesehen von der Inkubationstemperatur, die von 30 °C auf 37 °C erhöht wurde (Abschnitt 2.6.3).

2.4.1.3 Vorselektion einer Aminoacyl-tRNA Synthetasen-Bibliothek auf Agarplatten

Die Expressionsplasmide basierend auf pSB16.6c2, kodierten neben dem Ampicillin-Resistenzgen für das Chloramphenicol-Resistenz vermittelnde Gen der Chloramphenicol-Acetyltransferase, das ein Amber-Stop Codon an Position 112 trug. Dies ermöglichte die Selektion von Aminoacyl-tRNA Synthetasen-Varianten der Bibliothek auf die Fähigkeit, die nicht-natürlichen Aminosäuren zu erkennen und am Amber-Stop Codon in das Resistenzprotein eingebaut wurde.

Für die Etablierung der Vorselektion der Aminoacyl-tRNA Synthetasen wurden Agarplatten gegossen, die zusätzlich zu 100 mg/L Ampicillin und 1 mM der nicht-natürlichen Aminosäure (D/L-Bpa) bei auch Chloramphenicol mit einer Endkonzentration von 10 mg/L, 15 mg/L, 20 mg/L, oder 25 mg/L enthielten. Nach Transformation von pSB16.6c2 in elektrokompente *E. coli* Zellen (Abschnitt 2.2.3) und anschließender Regeneration in SOC-Medium wurde eine Verdünnungsreihe der Zellsuspension angefertigt und diese auf den vorbereiteten Selektionsplatten ausplattiert. Die Inkubation erfolgte für 22 h bei 37 °C, wonach die Transformanzahl bestimmt wurde und die adäquate Chloramphenicol-Konzentration für die Vorselektion der Aminoacyl-tRNA Synthetase-Bibliothek gewählt wurde. Zum Vergleich wurden die Transformanden auch auf LB-Agarplatten (100 mg/L Ampicillin) ausplattiert, ohne den Selektionsdruck durch Chloramphenicol. Die Vorselektion der Aminoacyl-tRNA Synthetase-Bibliothek erfolgte analog dazu mit der zuvor evaluierten Chloramphenicol-Konzentration und mit 1mM L-Hca.

2.5 Proteinchemische Methoden

2.5.1 Affinitätschromatographie an Streptavidin-Sepharose

Die Produktion rekombinanter Proteine, welche am C-Terminus mit dem *Strep*-tag II Affinitätsanhängsel fusioniert waren, ermöglichte es, diese Proteine aus der periplasmatischen Zellfraktion einer *E. coli*-Kultur mit Hilfe der Streptavidin-Affinitätschromatographie zu isolieren (Schmidt & Skerra, 1994; Schmidt & Skerra, 2007).

Der Periplasmaextrakt (Abschnitt 2.4.1.1) wurde dazu über Nacht bei 4 °C gegen das etwa 100fache Volumen an SA-Puffer (100 mM Tris/HCl, 500 mM Saccharose, 1 mM EDTA, pH 8,0) dialysiert und vor dem Auftragen auf die Chromatographie-Säule zur Entfernung von Proteinaggregaten sterilfiltriert (0,45 µm). Aus 2 L Flüssigkultur resultierte gewöhnlich ein Volumen von 30 ml dialysiertem Periplasmaextrakt, der mit Hilfe einer Pumpe auf eine mit SA-Puffer äquilibrierte Strep-Tactin-Chromatographiesäule aufgetragen wurde. Das Bettvolumen dieser Chromatographiesäule betrug in der Regel 2 ml. Die Säule wurde bis zum Erreichen einer konstanten Basislinie, detektiert anhand der Absorption bei 280 nm (A_{280}), mit SA-Puffer bei konstanter Flussrate gespült. Das über das *Strep*-Tag II gebundene rekombinante Protein wurde daraufhin mit einer 2,5 mM Lösung an D-Desthiobiotin in SA-Puffer eluiert und diese Fraktion anschließend hinsichtlich der Reinheit durch SDS-PAGE überprüft (Abschnitt 2.5.5). Die Säule wurde durch Spülen mit 5 mM HABA in SA-Puffer bis zu einer intensiven Orangefärbung regeneriert und im Anschluss wurde das gebundene HABA durch Spülen mit SA-Puffer bis zur vollständigen Entfärbung der Säule ausgewaschen. Bei Proteinproduktionen im größeren Maßstab wurden die dialysierten und filtrierten Periplasmaextrakte bis zu einem Gesamtvolumen von 150 ml vereint und auf eine Säule mit einem Bettvolumen von 12 ml aufgetragen.

2.5.2 Metallchelate-Affinitätschromatographie

Wie in Abschnitt 3.1.4 angegeben wurden einige Borocalin-Varianten mit einem C-terminalen Hexahistidinanhängsel (His₆-tag) in *E. coli* BL21 produziert. Daher wurden diese Varianten nach dem Periplasmaaufschluss mittels Ni²⁺-vermittelter Affinitätschromatographie (*Immobilized Metal Affinity Chromatography*, IMAC, (Porath *et al.*, 1975)) aus dem zellfreien Proteinextrakt isoliert (Skerra *et al.*, 1991).

Der Periplasmaextrakt (Abschnitt 2.4.1.1) wurde hierfür über Nacht bei 4 °C gegen das etwa 100fache Volumen an IMAC-Laufpuffer (0,5 M NaCl, 40 mM NaH₂PO₄, pH 7,4) dialysiert. Vor dem Auftragen auf die mit 0,1 M NiSO₄ beladene und ebenfalls mit IMAC-Laufpuffer äquilibrierte Chromatographie-Säule (Ni-Sepharose High Performance Säule, 1 ml Bettvolumen) wurde der Periplasmaextrakt zur Entfernung von Proteinaggregaten sterilfiltriert (0,45 µm). Die Flussrate der gesamten Chromatographie an der Äkta FPLC-Anlage betrug

dabei 4 ml/min. Nach dem Auftragen der Proteinprobe wurde die Säule bis zum Erreichen einer konstanten Basislinie (A_{280}) mit dem IMAC-Laufpuffer bei konstanter Flussrate gespült. Die Elution der an die Chromatographie-Säule gebundenen Proteine wurde mit einem zweistufigen Imidazol-Konzentrationsgradienten (0 bis 150 mM und 150 bis 300 mM Imidazol im Chromatographiepuffer) eluiert. Das Eluat wurde sofort nach der Fraktionierung mit 1 mM EDTA versetzt. Die Fraktionen des Haupt-*Peaks* welche das Zielprotein beinhalten, wurden vereint und mit Hilfe einer SDS-PAGE auf Reinheit geprüft. Die Säule wurde im Anschluss mit jeweils 10 Säulenvolumen ddH₂O, Stripping Buffer (1 M NaCl, 50 mM EDTA, pH 8), ddH₂O, 1 M NaOH, ddH₂O regeneriert und schließlich in 20 % EtOH gelagert.

2.5.3 Ionenaustauschchromatographie

Die finale Reinigung der Borocalin-Varianten erfolgte mittels Ionenaustauschchromatographie, bei der die Proteine aufgrund ihrer spezifischen pH-abhängigen Oberflächenladung aus einem Proteingemisch getrennt wurden. Für die reversible Bindung von positiv geladenen Seitenketten der Proteine an die negativ geladenen Methylsulfonat-Gruppen wurde bei dieser Arbeit eine Resource S-Kationenaustauschersäule (1 ml Bettvolumen) verwendet. Die zuvor gegen das 100fache Volumen an CEX-Laufpuffer (20 mM MES, pH 6) dialysierte und sterilfiltrierte (0,45 μ m) Proteinprobe wurde auf die ebenfalls mit CEX-Laufpuffer äquilibrierte Säule geladen. Die Flussrate der gesamten Chromatographie an der Äkta FPLC-Anlage betrug dabei maximal 4 ml/min. Die Säule wurde nach dem Auftrag der Proteinprobe bis zum Erreichen der Basislinie (A_{280}) mit CEX-Laufpuffer gespült und durch Anlegen eines linearen Konzentrationsgradienten bis 500 mM NaCl in CEX-Laufpuffer fraktioniert eluiert. Die Fraktionen des Haupt-*Peaks* wurden vereint und Proben für die Proteinauftrennung mittels SDS-PAGE (Abschnitt 2.5.5) entnommen.

2.5.4 Größenausschlusschromatographie

In dieser Arbeit wurde die Größenausschlusschromatographie (*Size-Exclusion Chromatography*, SEC) als Feinreinigung der Borocalin-Variante 36-Bpa(NFW) nach der Ionenaustauschchromatographie (Abschnitt 2.5.3) vor der Protein-Kristallisation verwendet. Die Trennung der Proteine erfolgt hierbei anhand ihrer Größe sowie Form mit dem porösen Säulenmaterial. Kleinere Moleküle haben ein größeres zugängliches Diffusionsvolumen als größere und durch den so auftretenden molekularen Siebeffekt kommt es zur Trennung, die sich umgekehrt proportional zum Logarithmus des apparenten Molekulargewichts der Proteine verhält. Die zuvor gegen reichlich PBS dialysiert und mit Ultrafiltrationsröhrchen auf 500 μ l eingengte Proteinprobe wurde auf eine ebenfalls zuvor mit PBS äquilibrierte HiLoad 16/60 Superdex S75 Säule (Bettvolumen 120,6 ml) geladen. Die Chromatographie wurde bei einer konstanten Flussrate von 0,5 ml/min mit PBS als Laufpuffer betrieben. Die gesammelten Elutionsfraktionen wurden durch SDS-PAGE (Abschnitt 2.5.5) analysiert.

2.5.5 SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE)

Für analytische Zwecke diente die größenabhängige Trennung von Proteingemischen unter denaturierenden Bedingungen durch diskontinuierliche SDS-PAGE (Laemmli, 1970) unter Verwendung eines Hochsals-Puffersystems (Fling & Gregerson, 1986). Die Methode beruht auf der Denaturierung der Proteine mit dem Detergens Natriumdodecylsulfat (SDS) und der damit verbundenen hydrophoben Anlagerungen, welche in einer negativen Gesamtladung proportional zur Molekülgröße resultieren. Hinzu kommt der Molekularsiebeffekt des Polyacrylamid-Gels, wodurch sich die elektrophoretische Mobilität der Proteine annähernd umgekehrt proportional zum Logarithmus der Molmasse verhält. Aufgrund der erwarteten molekularen Masse der produzierten rekombinanten Proteine wurden in dieser Arbeit immer 15 %-ige SDS-Polyacrylamidtrenngel und 5 %-ige Sammelgele verwendet.

Für die Herstellung der Acrylamid-Gele wurden die Glasplatten gereinigt, in die Apparatur eingespannt und der Zwischenraum der Platten mit 4,5 ml der Trenngelmischung befüllt, welche vorsichtig mit 500 ml ddH₂O überschichtet wurde. Nach etwa 40 min konnte das ddH₂O von dem auspolymerisierten Trenngel entfernt und die Sammelgelmischung in die Kammer gefüllt werden. Die Geltaschen wurden durch das Einbringen eines Kammes in die noch flüssige Sammelgelmischung geformt. Nach weiteren 40 min konnte der Taschenkamm aus dem polymerisierten Sammelgel entfernt und die mit dem Acrylamid-Gel befüllten Glasplatten aus der Apparatur entnommen werden. Die so vorgefertigten Gele wurden bis zu 4 Wochen im Kühlschrank gelagert. Vor der PAGE wurden die Proben in einem Verhältnis von 4 zu 1 mit 5x Auftragspuffer versetzt, für 5 min auf 95 °C erhitzt und die mit Acrylamid-Gel befüllten Glasplatten in die mit 1x SDS-Laufpuffer gefüllte Elektrophoresekammer eingespannt. Im Anschluss daran wurde typischerweise jeweils 10 µl der Ganzzellproben und 5 µl der Proteinproben, die nach dem Zellaufschluss oder einem Reinigungsschritt entnommen wurden, in die Geltaschen geladen. Um die Proteingröße abschätzen zu können, wurde auch ein Protein-Größenstandard aufgetragen. Die Elektrophorese wurde bei 200 V durchgeführt, bis der Bromphenolblau-Marker die untere Gelkante erreichte. Nach Entfernen der Glasplatten wurden die aufgetrennten Proteine im Acrylamid-Gel durch 30-minütiges Schwenken in einer Commassie-Brilliantblau-Färbelösung, gefolgt von Entfärbelösung 1 für 15 min und Entfärbelösung 2 über Nacht, sichtbar gemacht.

2.5.6 Konzentrierung von Proteinen

Gegebenenfalls wurde die Konzentration von Proteinlösungen erhöht unter Verwendung von Zentrifugen-Filtereinheiten mit einer Ausschlussgröße von 10 kDa. Vorab wurde die Konzentratormembran mit 5 ml des verwendeten Puffers gespült und daraufhin die Proteinlösung durch Zentrifugation (2500 rpm, 4 °C) konzentriert. Die Proteinlösung wurde

anschließend in ein frisches Reaktionsgefäß überführt und die Konzentration durch Messung der Absorption bei 280 nm bestimmt.

2.5.7 Bestimmung von Proteinkonzentrationen

Die Quantifizierung der Proteinkonzentration einer Lösung erfolgte mittels Absorptionsspektrometrie (A_{280}), basierend auf dem Lambert-Beerschen Gesetz ($A = \epsilon \cdot c \cdot d$). Der hierfür notwendige molare Extinktionskoeffizient (ϵ) wurden anhand der Summe der Absorptionsbeiträge aller in der Polypeptidkette enthaltenen Tyr-, Trp- und Cys-Seitenketten berechnet (Gill & von Hippel, 1989), wobei die nicht-natürliche Aminosäure Bpa mit Tyr gleichgesetzt wurde. Für die Berechnung wurde das ProtParam-Tool des ExPASy-Servers (www.expasy.org) verwendet (Gasteiger et al., 2003).

2.6 Biophysikalische Methoden

2.6.1 Fluoreszenzspektroskopie

Die Bindungseigenschaften der Borocalin-Varianten zu Liganden wurden mittels Fluoreszenztitration bestimmt (Breustedt *et al.*, 2006). Gemessen wurde dabei die Zu- oder Abnahme der intrinsischen Trp- und Tyr-Fluoreszenz aufgrund der durch die Komplexbildung veränderten Umgebung. Die Messungen wurden mit einem Perkin Elmer Fluoreszenzphotometer (LS 50 B) durchgeführt. Dafür wurde eine Anregungswellenlänge von 294 nm und eine Emissionswellenlänge von 345 nm, jeweils mit einer Spaltbreite von 4,5 nm, gewählt.

Für die Messung wurden 2 ml einer 1 μ M Borocalin-Lösung (50 mM CHES/NaOH, 1 mM EDTA, pH 8,5) in einer Quarzküvette (1 cm²) mit einem Rührfisch vorgelegt und für 15 min im Probenhalter des Photometers auf 20 °C temperiert. Im Anschluss wurden bis zu 100 μ l der entsprechenden Liganden-Stammlösung in Schritten von 2 μ l bis 10 μ l hinzugefügt. Die dabei stattfindende Verdünnung belief sich auf maximal 5 % und wurde vernachlässigt. Nach jedem Titrationsschritt wurde zur Gleichgewichtseinstellung bei verdecktem Anregungslicht für 1 min unter Rühren inkubiert, bevor das Fluoreszenzsignal während einer Integrationszeit von 5 s gemessen wurde. Die gemessenen Werte wurden um den inneren Filtereffekt des Liganden wie unten beschrieben korrigiert und auf einen Anfangswert von 100 % normiert.

Obwohl die Liganden teilweise nur geringe Eigenabsorption aufwiesen, wurde für alle Liganden der Korrekturfaktor berechnet und die Messwerte dahingehend korrigiert (Breustedt *et al.*, 2006). Die Abnahme der gemessenen Fluoreszenzintensität in Abhängigkeit von der eingesetzten Konzentration des jeweiligen Liganden wurde durch Fluoreszenztitration mit einer 5 μ M N-Acetyl-L-tryptophanamid (in 50 mM CHES/NaOH, 1 mM EDTA, pH 8,5) bestimmt. Da die verwendeten Liganden mit N-Acetyl-L-tryptophanamid nicht durch molekulare

Komplexbildung wechselwirken, war die Abnahme der Fluoreszenzintensität auf deren Eigenabsorption zurückzuführen.

Die Korrektur der jeweils gemessenen Fluoreszenzintensität (F_{App}) bei der Komplexbildung zwischen Borocalin (P) und Ligand (L) wurde unter Anwendung des Lambert-Beerschen Gesetzes mit folgender Formel berechnet:

$$F = F_{App} \cdot 10^{\varepsilon_{App} c L} \quad (1)$$

Im Verlauf der Fluoreszenztitration lässt sich die Komplexbildung aus Protein und Ligand durch das Massenwirkungsgesetz beschreiben. Die Dissoziationskonstante K_D wurde durch die Kurvenanpassung der Daten mittels nicht-lineare Regression unter Verwendung des Computerprogrammes KaleidaGraph ermittelt. Folgende Formel wurde dafür verwendet:

$$F = ([P]_t - [L]_t - K_D) \frac{f_P}{2} + ([L]_t - [P]_t - K_D) \frac{f_L}{2} + ([P]_t + [L]_t + K_D) \frac{f_{PL}}{2} \quad (2)$$

$$+ (f_P + f_L - f_{PL}) \sqrt{\frac{([P]_t + [L]_t + K_D)^2}{4} - [P]_t [L]_t}$$

Dabei stehen $[P]_t$ und $[L]_t$ für die totale Konzentration der Proteinlösung (konstant) beziehungsweise der zugegebenen Ligandenlösung nach jedem Titrationschritt. f_P , f_L und f_{PL} beschreiben den relativen molaren Extinktionskoeffizient des freien Proteins, des freien Liganden oder des Protein/Ligand-Komplexes. K_D und f_{PL} wurden dabei als freie Parameter zur Kurvenanpassung verwendet (Breustedt *et al.*, 2006).

2.6.2 Massenspektrometrie

Die exakte Bestimmung der Molekülmasse von den in *E. coli* produzierten rekombinanten Proteinen, gegeben Falls unter Einbau einer nicht-natürlichen Aminosäure, wurde mit Hilfe eines ESI-Q-TOF-Massenspektrometers des Typs MaXis (Bruker) durchgeführt.

Ein Aliquot der Proteinlösung (50 μ l, mindestens 5 μ M) wurde dreimal gegen 1 L 10 mM Ammoniumacetat (pH 6,8) für mindestens 4 h in Mikrodialyseeinheiten (*Slide-A-Lyzer Dialysis Unit*, MWCO 10 kDa, Thermo Fisher Scientific) dialysiert. Kurz vor der Messung wurde 20 % (v/v) Acetonitril und 0,1 % (v/v) Ameisensäure zugegeben. Die Messung wurde so unter denaturierenden Bedingungen im Positiv-Ionenmodus durchgeführt. Die daraus resultierenden Massenspektren wurden mit der Software Data Analysis 4.3 von Bruker mit Hilfe des MaxEnt Algorithmus (Ferrige *et al.*, 1991) über den gewünschten Massenbereich dekonvolviert und ausgewertet.

Um auch Protein/Ligand-Wechselwirkungen mittels Massenspektrometrie nachweisen zu können, wurden Messungen unter quasi „nativen“ Bedingungen vorgenommen (Leney & Heck, 2017). Dafür wurde ein Aliquot der Proteinlösung dreimal gegen 0,1 mM

Ammoniumbicarbonat/NH₃ pH 8,85 dialysiert und die Probe für 15 min bei Raumtemperatur mit 1 mM der Ligandenlösung inkubiert. Die Ionentransferparameter und die Kalibrierung des Instruments bei m/z von 2,500 wurde mit dem *ESI Tuning Mix* der Firma Agilent optimiert. Um ein intensiveres Signal zu erhalten wurde die kollisionsinduzierte Dissoziationsenergie (CID) bei dieser Messung von 0 V auf 30 V und die Kollisionsspannung von 8 V auf 60 V erhöht (Edwardraja *et al.*, 2017).

2.6.3 Cytofluorimetrische Selektion aus Gen-Bibliotheken und Einzelklonanalyse mittels *Fluorescence Activated Cell Sorting* (FACS)

Fluorescence Activated Cell Sorting ermöglicht die quantitative Analyse und molekulare Charakterisierung von Zellen in einem Hochdurchsatz-Verfahren. Dabei wird eine Zellsuspension mittels hydrodynamischer Fokussierung mit bis zu 20.000 *Events/sec* an einem monochromatischen Laserstrahl vorbeigeführt. Nach dem Passieren des Laserstrahls kommt es zum Tropfenabriss, wobei die Tröpfchengröße so gewählt wird, dass sich statistisch nur eine Zelle darin befindet. Die Form, Struktur oder Färbung der Zelle erzeugt durch den Laserstrahl ausgelöste Signale in Form von Lichtstreuung oder Fluoreszenzemission. Die Lichtstreuung gibt Aufschluss über das Volumen (*Forward Scatter*, FSC) und die Granularität der Zelle (*Side Scatter*, SSC). Da Zellen üblicherweise kein Fluoreszenzlicht emittieren, kann man mit Hilfe von Fluoreszenzfarbstoffen Protein/Liganden-Bindung (Daugherty *et al.*, 2000) oder, wie es in dieser Arbeit der Fall war, die Expression eines fluoreszierenden Reporterproteins detektieren (Kuhn *et al.*, 2010). Die Ergebnisse der tausendfachen Messungen werden in *Dot-Plots* oder als Histogramme dargestellt. Ersteres ermöglicht so eine Assoziation von zwei Parametern, die zur Auswahl der zu sortierenden Zellen herangezogen werden. Die Sortierung der Zellen beruht auf der Ladung, die jeder Tropfen beim Tropfenabriss erhält, welcher in der darauffolgenden Fallstrecke beim Passieren des elektrischen Feldes eines Plattenkondensators abgelenkt werden kann.

Die *E. coli* Zellen aus 1 ml-Proben nach 3-stündiger oder 16-stündiger Genexpression des Reporterproteins eGFP(39a) in einem Kulturvolumen von 5 ml (Abschnitt 2.4.1.2) wurden in Röhren bei 2000 g, 4°C für 10 min (Sigma 4-16K) sedimentiert. Das Zellsediment wurde mit 1 ml PBS gewaschen, erneut zentrifugiert und in 1 ml PBS resuspendiert. Für die Analyse im FACS-Gerät sowie für den Sortiervorgang wurden die in PBS resuspendierten Zellen im Verhältnis von 1:50 mit PBS verdünnt und in ein FACS-Messröhrchen (5 ml) überführt. Die Verdünnung führte dazu, dass durch die Anpassung der Flussgeschwindigkeit die Messereignisdichte (Event Rate) auf 10.000 *Events/s* (Analyse) beziehungsweise 16.000 *Events/s* (Sortier-Vorgang) reguliert werden konnte.

Während der gesamten Probenvorbereitung wurden die Zellen auf Eis gelagert. Bevor die Messung der einzelnen Proben gestartet wurde, wurden die Fluoreszenzdetektoren des Geräts

mit CS&T-Partikeln kalibriert. Die Analyse der durch die eGFP(39a)-Expression erzeugten Fluoreszenz erfolgte durch Anregung mit einem Laserstrahl bei 488 nm und Detektion mit dem FITC-Emissionsfilter (530/30 nm). Für ein statistisch aussagekräftiges Ergebnis der Analyse wurden die Fluoreszenzsignale von 100.000 Zellen gemessen und in weiterer Folge für die Einstellung des Sortier-Fensters herangezogen. Vor jedem Sortier-Vorgang wurde die Distanz zwischen dem Fluoreszenzsignal in der Messküvette des Gerätes und dem Tropfenabriss des Flüssigkeitsstrahls mit Accudrop-Partikeln ermittelt und entsprechend der Bedienungsanleitung des Herstellers zur Kalibrierung des FACS-Systems verwendet.

Die anhand ihres Fluoreszenzsignals ausgewählten Zellen wurden in ein 15 ml Reaktionsgefäß, in dem 2 ml LB-Medium vorgelegt war, sortiert. Jeweils 2 ml der erhaltenen Zellsuspension wurde im Anschluss auf quadratische Agarplatten (100 mg/L Ampicillin oder 30 mg/L Chloamphenicol) ausplattiert und über Nacht bei 37 °C inkubiert. Auf die dicht bewachsene Agarplatte, auf der gerade noch Einzelkolonien sichtbar waren, wurde 10 ml LB-Medium mit den 100 µg/ml Ampicillin pipettiert, die Zellen von der Platte geschabt und diese durch 30-minütiges schütteln vollständig resuspendiert. Aus dieser Zellsuspension wurde eine neue Flüssigkultur für einen weiteren Selektionszyklus inokuliert oder für die Präparation der Plasmid-DNA, gefolgt von einer Zufallsmutagenese mittels *Error-Prone* PCR zur Diversifizierung der Bibliothek, verwendet. Alternativ konnten auch Kolonien von den Platten gepickt werden, um eine Einzelklonanalyse zu ermöglichen. Zur Auswertung der Cytofluorometrie-Daten wurde die FlowJo-Software benutzt.

2.7 Proteinkristallographische Methoden

Für die strukturelle Untersuchung der Borocalin-Variante 36Bpa-NFW hinsichtlich der Bildung eines zyklischen Esters zwischen der Boronsäure und 4-Nitrocatechol wurde 36Bpa-NFW im Komplex mit 4-Nitrocatechol von Dr. Andreas Eichinger am Lehrstuhl für Biologische Chemie kristallisiert. Die initialen Kristallisationsbedingungen wurden mit einem am Lehrstuhl entwickelten *Sparse-Matrix Screen* (Jancarik & Kim, 1991) identifiziert. Dazu wurde der aus 480 Kristallisationsbedingungen bestehende *Vector Screen* verwendet, der am Lehrstuhl für Biologische Chemie zusammengestellt wurde. Die Kristallisationsversuche wurden nach der Dampfdiffusionsmethode mit sitzenden Tropfen in 96-Well Crystal Quick-Platten (Thermo Fisher Scientific) durchgeführt. 4-Nitrocatechol (in ddH₂O) wurde einer konzentrierten Proteinlösung (in 20 mM CHES, pH 8,5) zugesetzt, davon 400 nl mit Hilfe des Roboters Freedom Evo im Verhältnis 1:1 (v/v) zu der im gleichen Volumen vorgelegten Reservoirlösung pipettiert und gegen 100 µl Reservoirlösung äquilibriert. Die Platten wurden nach dem Verschluss mit transparenter Klebefolie bei 20 °C inkubiert und der Kristallisationsverlauf mit Hilfe des Rock Imager Systems verfolgt.

Vielversprechende Kristallisationsbedingungen aus dem *Vector Screen* wurden im *Fine Screen* hinsichtlich Größe und Streufähigkeit der entstandenen Kristalle durch Veränderung der Parameter (Proteinkonzentration, Fällungsreagenz, pH-Wert, Pufferkonzentration) optimiert. Dies wurde in 24-Well Platten mit der Dampfdiffusionsmethode im hängenden Tropfen gegen 1 ml der entsprechenden Reservoirlösung erreicht. Dafür wurde der Komplex aus Protein und Ligand im Verhältnis 1:1 mit der Reservoirlösung auf silanisierten Deckgläschen gemischt (jeweils 400 nl), diese umgedreht auf dem Reservoir platziert, mit einer Öl/Vaseline-Mischung abgedichtet und bei 20 °C inkubiert. Die gezüchteten Kristalle wurden geerntet, nach Zugabe von 30 % (v/v) Glycerol eingefroren und von Dr. Andreas Eichinger an der Synchrotronstrahlenquelle 14.1 des BESSY in Berlin vermessen. Aus den Beugungsdaten wurde die Kristallstruktur durch molekularen Ersatz unter Verwendung von Apo-36Bpa-NFW (PDB ID: 5MHH, Edwardraja *et al.*, 2017) als Suchmodell bearbeitet. Die Atomkoordinaten und Strukturaktoren der so gelösten Kristallstruktur sind in der Protein Data Bank (PDB) unter dem Identifikationskode 6QMU hinterlegt.

2.8 Verwendete Computerprogramme und Datenbanken

A plasmid Editor 2.0.61 (ApE)	Wayne Davis, Salt Lake City, UT, USA
Adobe Illustrator CC	Design Standard Adobe Systems, San Jose, CA, USA
ChemDraw Ultra 12.0	PerkinElmer, Rodgau
Data Analysis Software 4.0	Bruker Corporation, Billerica, MA, USA
EndNote X9	Clarivate Analytics, Philadelphia, PA, USA
ExpASy; ProtParam	(Gasteiger <i>et al.</i> , 2003) (http://web.expasy.org/protparam)
GraphPad PRISM 6.01	GraphPad Software, San Diego, CA, USA
NanoDrop 2000/2000c 1.6.198	Thermo Fisher Scientific, St. Leon-Rot
Office 365 ProPlus	Microsoft, Redmond, WA, USA
Protein Data Bank	(Berman <i>et al.</i> , 2000) (www.rcsb.org/pdb)
PyMol 2.2.2	(DeLano, 2002)
SnapGene 3.3	GSL Biotech LLC, San Diego, CA, USA
Unicorn 5.20	GE Healthcare, Freiburg
KaleidaGraph 4.1	Synergy Softwar, Reading PA, USA
FlowJo (10.5.3)	Tree Star, Ashland, USA
FACSDiva (5.0.2)	Becton Dickinson, New Jersey, USA

CRISPR-Tool

<http://crispor.tefor.net>

(Concordet & Haeussler, 2018)

Rock Imager[®]

Formulatrix, Bedford, Massachusetts,
USA

Freedom EVO[®]

Tecan, Männedorf, Schweiz

3 Ergebnisse

3.1 Engineering eines Lipocalins mit einer nicht-natürlichen Aminosäure für die Bindung von Zuckermolekülen

Wesentliche Vorarbeiten zur Funktionalisierung eines Lipocalins mit einer nicht-natürlichen Aminosäure, um eine selektive und hochaffine Zuckerbindung zu ermöglichen, wurden von Dr. Selvakumar Edwardraja in der Arbeitsgruppe von Prof. Dr. Skerra durchgeführt (Edwardraja *et al.*, 2017). Anhand der Kristallstruktur des humanen Lcn2 (PDB ID: 1L6M) wurde eine geeignete Position für den translationalen Einbau der nicht-natürlichen Aminosäure Bpa in der Bindetasche des Lipocalins gesucht.

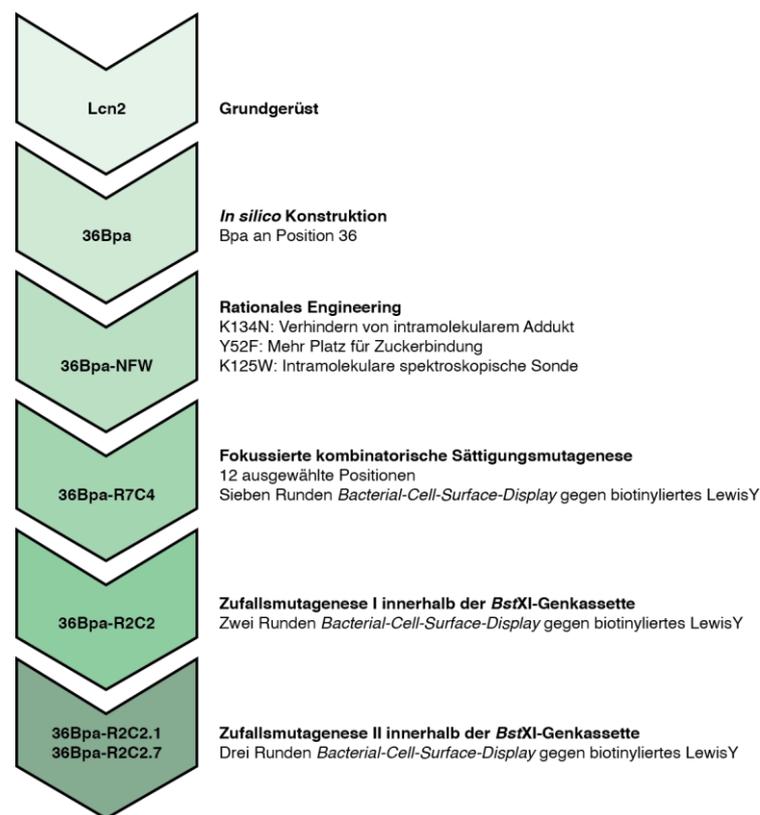


Abbildung 7: Historie der von Dr. Selvakumar Edwardraja generierten Borocalin-Varianten. Zunächst wurde *in silico* eine geeignete Position für Bpa identifiziert sowie die Bindetasche rational modifiziert. Im Anschluss folgte eine gerichtete Evolution. Ausgehend von der Variante 36Bpa-NFW wurde eine positionsbezogene Sättigungsmutagenese mittels NNS-Codons durchgeführt und auf biotinyliertes LewisY selektiert. Darauf folgten zwei Zufallsmutagenesen in Verbindung mit zwei Selektionskampagnen. Die Primärsequenz des Lcn2 Grundgerüsts sowie die verwendenden Varianten ist in Abbildung 8 dargestellt.

Bei dem hierfür erforderlichen molekularen Modelling wurde darauf geachtet, dass Bpa für eine Zuckerbindung frei zugänglich ist und dessen Einbau zu keinen sterischen Konflikten mit benachbarten Seitenketten führt. Diese Kriterien wurden für die nicht-natürliche Aminosäure Bpa an der Aminosäure-Position 36 erfüllt. Der ortsspezifische ribosomale Einbau von Bpa

wurde mit Hilfe der Amber-Stop Codon Suppressions-Methode realisiert. Hierfür wurde das für die Aminosäure (Leu) an Position 36 kodierende Codon zu einem Amber-Stop Codon mutiert. Das für die Unterdrückung der Transkriptionstermination am Amber-Stop-Codon verwendete orthogonale Paar aus Aminoacyl-tRNA Synthetase und Suppressor-tRNA umfasst eine für die molekulare Erkennung von Bpa modifizierte Tyrosyl-tRNA Synthetase aus *Methanocaldococcus janaschii* (*MjTyrRS*), genannt BpaRS (Brustad *et al.*, 2008), und eine für die *MjTyrRS* optimierte Suppressor-tRNA (Wang *et al.*, 2001).

Zusätzlich zu der Einführung der Bpa-Seitenkette in Lcn2 wurde Lys¹³⁴ zu Asn mutiert, um eine intramolekulare Adduktbildung zwischen der Lysin-Seitenkette und der Boronsäure zu verhindern. Des Weiteren wurde Tyr⁵² zu Phe mutiert, um für eine Zuckerbindung ausreichend Platz zu schaffen. Als intramolekulare spektroskopische Sonde für die Detektion der Liganden-Bindung wurde schließlich Lys¹²⁵ zu Trp mutiert (Edwardraja *et al.*, 2017). Die daraus resultierende sogenannte Borocalin-Variante wurde zur Kennzeichnung der vorgenommenen Mutationen als 36Bpa-NFW bezeichnet und diente als Ausgangsvariante für eine gerichtete Evolution zur selektiven Erkennung des Tumor-Zucker-Markers LewisY (Abbildung 7).

Dieser Prozess umfasste ein Wechselspiel aus Mutagenese und Selektionskampagnen, die von Dr. Selvakumar Edwardraja durchgeführt wurden. Ausgehend von einer Zufalls-Bibliothek, erzeugte Dr. Selvakumar Edwardraja eine fokussierte, kombinatorische Sättigungsmutagenese worauf zwei weiteren Zufallsmutagenesen mittels *epPCR* innerhalb der gesamten *BstXI*-Genkassette folgten. Die drei aus der Selektion gegen das biotinylierte Tetrasaccharid LewisY resultierenden Borocalin-Varianten, 36Bpa-R2C2, 36Bpa-R2C2.1 und 36Bpa-R2C2.7 (Abbildung 8), wurden neben 36Bpa-NFW am Beginn dieser Arbeit präparativ produziert und hinsichtlich ihrer Bindungseigenschaften untersucht.

	1	10	20	30	40	50	60	70	80	90
Lcn2	QDSTSDLI PAPP LSKV PLQQNFQDNQ FQ GKWYV VGLAGN AILREDK DPQKMYATI YELKEDKSYNVT SVLFRKKKCDY WIRTFVPGCQPGEFTLGNIKS									
36Bpa-NFWh.....*.....F.....s.....									
Sättigungsmutageneseh.....*.....xxx x xxx F.....x xxx.....s.....									
36Bpa-R2C2h.....*.....DNQ T A.GV F.....D.WHG.....s.....R.....									
36Bpa-R2C2.1h.....*.....DNQ T A.GV F.....D.WHG.ET.....s.....AR.Y.....									
36Bpa-R2C2.7h.....*.....DNQ T A.GV F.....D.WHG.....D.s.....R.D.....									
	====A===== #1-----B==== ===C==== #2-----D==== ==E---- #3									
	100	110	120	130	140	150	160	170	178	
Lcn2	YPGLT SYLVRVVSTNYNQHAMVFFKVSQNR EYFKITLYGR TKELTSELKENFIRFSKSLGLPENHIVFPVPIDQCIDG									
36Bpa-NFWW.....N.....									
SättigungsmutageneseW.....N.....									
36Bpa-R2C2W.....N.....									
36Bpa-R2C2.1	...I.G.V.....W.....N.....									
36Bpa-R2C2.7	...I.R.....W.....N.....									
	#3-----F===== ===G----- #4-----H=====									

Abbildung 8: Aminosäuresequenzen der am Beginn dieser Arbeit vorliegenden Borocalin-Varianten im Vergleich zu Lcn2. Die acht konservierten β -Faltblattstränge (A-H) sowie die vier strukturell variable Schleifen (#1-#4) sind gekennzeichnet. Randomisierte Aminosäurepositionen bei der Sättigungsmutagenese sind mit dem Buchstaben x markiert. Kleinbuchstaben kennzeichnen Aminosäureaustausche zum Einführen einer zweiten *BstXI*-Restriktionsschnittstelle (Q28H) sowie zum Entfernen des freien Cysteinrestes (C87S) in Lcn2.

3.1.1 Produktion und Reinigung rekombinanter Borocalin-Varianten

Die Produktion der Borocalin-Varianten 36Bpa-NFW (pSB4.6c14 (NFW)), 36Bpa-R2C2 (pSB6.6c92 (Lcn2aR2.C2)), 36Bpa-R2C2.1 (pSB6.6c92 (Lcn2aR2.C2.1)) und 36Bpa-R2.C2.7 (pSB6.6c92 (Lcn2aR2.C2.1)) erfolgte im Periplasma von *E. coli* BL21 als Fusion mit der OmpA-Signalsequenz (Movva *et al.*, 1980) und einem C-terminalen *Strep*-tagII (Schmidt & Skerra, 2007) unter der Kontrolle des lacUV5 Promotor/lac-Operator-Systems mit dem Expressionsplasmid pSB4.6c14 (Plasmid von Dr. Selvakumar Edwardraja zur Verfügung gestellt; Abbildung 9). Auf diesem Plasmid ist auch das orthogonale Paar, bestehend aus der BpaRS (Brustad *et al.*, 2008) unter der Kontrolle des tetA Promotor/Operator-Systems (Skerra, 1994) und der Suppressor-tRNA unter Kontrolle des konstitutiven lpp-Promotors (Nakamura & Inouye, 1979), kodiert.

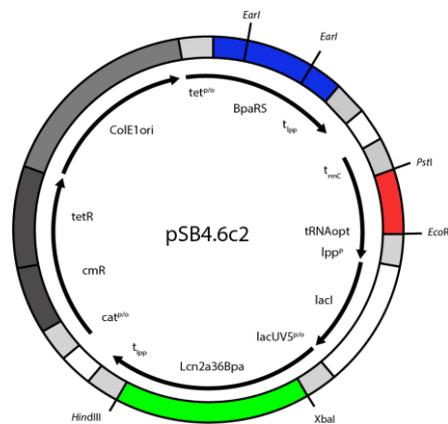


Abbildung 9: Schematische Darstellung des Expressionsplasmids für die Produktion eines Borocalins. Das Plasmid pSB4.6c14 kodiert für ein Fusionsprotein aus OmpA-Siganlsequenz (Movva *et al.*, 1980) und Borocalin inklusive des *Strep*-tagII (Schmidt & Skerra, 2007) am 3'-Ende unter der Kontrolle des mit IPTG induzierbaren lacUV5 Promotor/Operators (lacUV5^P) (Silverstone *et al.*, 1970), gefolgt von dem lpp-Terminator (t_{lpp}). Zusätzlich trägt das Plasmid die Gensequenz für die BpaRS (Brustad *et al.*, 2008) unter der Kontrolle des mit aTc regulierbaren tetA Promotor/Operators (tet^P) (Skerra, 1994), gefolgt vom lpp-Terminator (t_{lpp}) und der dazugehörigen Suppressor-tRNA (tRNA^{opt}) unter der Kontrolle des lpp-Promotors (lpp^P), terminiert vom rrnC-Terminator (t_{rrnC}) terminiert wird. Weitere Elemente auf dem Plasmid sind der ColE1 Replikationsursprung (ColE1 ori) und der Lac-Repressor (lacI) sowie dem Tet-Repressor (tetR). Die konstitutive Expression der Transkriptionsfusion für den Tet-Repressor und die Chloramphenicol-Acetyltransferase (camR) ermöglicht die effektive Repression des Promotors für die Expression der BpaRS wie auch die antibiotische Selektion.

Die bakterielle Kultur wurde nach der Anzucht über Nacht bei 30 °C im 2 L-Kulturmaßstab mit LB-Medium unter schütteln inkubiert. Die Zugabe der nicht-natürlichen Aminosäure D/L-Bpa erfolgte bei OD₅₅₀ = 0,3 und vor der schrittweisen Induktion der BpaRS mit 50 ng/μl Anhydrotetracyclin (aTc) und der Borocalin-Varianten mit 500 mM Isopropyl-β-D-thiogalactopyranosid (IPTG) nach jeweils einer Stunde Inkubationszeit. Die weitere Genexpression erfolgte über Nacht (~ 16 h) und zum Zeitpunkt der Zellernte betrug die OD₅₅₀ ungefähr 3,2 bis 4,7 (Abschnitt 2.4.1.1). Aufgrund des Signalpeptids wurde das rekombinante Protein über die innere Membran der Bakterienzelle transloziert, wobei das Signalpeptid abgespalten wurde. Das ermöglichte die Reinigung des rekombinanten Proteins aus dem periplasmatischen Zellextrakt mittels *Strep*-tagII-Affinitätschromatographie (Abschnitt 2.5.1). Abschließend wurde das rekombinante Protein durch AEX (Abschnitt 2.5.3)

zur Homogenität gereinigt (Abbildung 10 A und B). Der gewünschte Einbau der nicht-natürlichen Aminosäure Bpa wurde mittels Massenspektrometrie nachgewiesen (Abbildung 10 C; Abschnitt 2.6.2), wozu ein Aliquot der Proteinlösung gegen Ammoniumacetat Puffer pH 6,6 dialysiert wurde. Die Proteinausbeuten aus zwei Liter Kultur betrugen für 36Bpa-NFW 1,4 - 1,8 mg, für 36Bpa-R2C2 ca. 1,5 mg für 36Bpa-R2C2.1 1,2 - 1,4 mg und für 36Bpa-R2C2.7 0,4 - 0,5 mg. Dagegen wurde im Fall der Expression und Reinigung des rekombinanten Proteins mit einem Tyrosin an Position 36 (36Y-NFW) anstelle von Bpa eine Ausbeute von 8,8 mg aus zwei Liter Kultur erzielt.

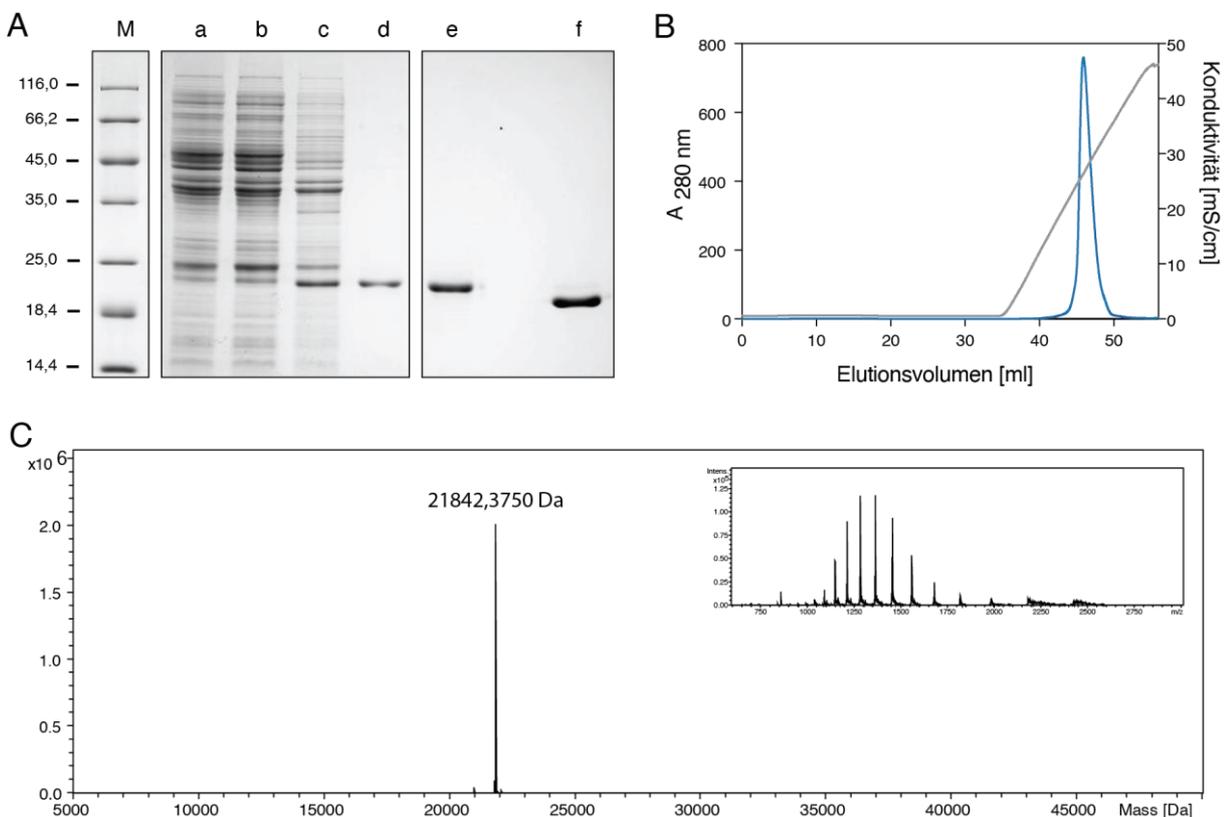


Abbildung 10: Überblick über die gentechnische Herstellung und Reinigung der Borocalin-Varianten sowie des Einbaus der nicht-natürlichen Aminosäure am Beispiel von 36Bpa-NFW. (A) 12 %-iges SDS-Gel, wovon bei den Proben in Spur a – e reduzierende Bedingungen und in Spur f nicht reduzierende Bedingungen vorliegen. Spur M: Proteingrößenstandard; Spur a: bakterieller Gesamtzellextrakt vor der Induktion der BpaRS und der Borocalin-Variante; Spur b: bakterieller Gesamtzellextrakt zum Zeitpunkt der Zellernte; Spur c: periplasmatische Proteinextrakt; Spur d: Elutionsfraktion der Streptavidin-Affinitätschromatographie; Spur e bzw. f: Elutionsfraktion der c. Das schnellere Laufverhalten (Spur f) ist auf die Disulfidbrücken im denaturiertem Protein unter nicht-reduzierenden Bedingungen zurückzuführen. (B) Abschließende Reinigung mittels Anionenaustauschchromatographie. (C) ESI-Massenspektrum des gereinigten Proteins Lcan2a36Bpa-NFW, $MW_{\text{kalk.}} = 21842,59 \text{ Da}$; $MW_{\text{obs.}} = 21842,38 \text{ Da}$; $\Delta m = 0,21 \text{ Da}$.

3.1.2 Evaluierung der Bindungsaffinität von Borocalin-Varianten gegenüber Zuckern und Zuckerderivaten

Die Affinität der Borocalin-Varianten zu Zuckermolekülen und Zuckerderivaten beruht auf der Eigenschaft der Borsäure, reversible zyklische Diester mit *cis*-Diolen zu bilden (Peters, 2014). Der Boronsäurerest von Bpa ($pK_S \sim 9$) reagiert bei hohem pH-Wert als Lewis Säure mit einer

Lewis Base wie zum Beispiel OH^- und bildet ein tetraedrisches Boronatanion (Pizer & Babcock, 1977). Die Affinität der Borocalin-Varianten zu Zuckermolekülen und Zuckerderivaten wurde mittels Fluoreszenztitration in einem CHES/ EDTA-Puffer mit einem pH-Wert von 8,5 bestimmt. Diese Methode beruht auf der Änderung der intrinsischen Proteinfluoreszenz, welche auf die aromatischen Aminosäuren Tryptophan und Tyrosin zurückzuführen ist. Verantwortlich ist die lokale Veränderung der Polarität in Folge einer Komplexbildung mit dem Liganden in konzentrationsabhängiger Weise. Die Trp-Seitenketten wurden selektiv bei 295 nm angeregt und ein Fluoreszenzsignal bei der Emissionswellenlänge von 345 nm detektiert. Aus der Veränderung des Fluoreszenzsignals konnte nach der Normierung der Daten die Dissoziationskonstante (K_D -Wert) mittels Kurvenanpassung (Breustedt *et al.*, 2006) berechnet werden konnte (Abschnitt 2.6.1).

Die Affinität von 36Bpa-NFW wurde gegenüber den Monosacchariden D-Galactose, D-Glucose und D-Mannose sowie den Modell-Diolen 4-Nitrocatechol und Sorbit (Abbildung 11) bestimmt. Hierbei erwies sich 4-Nitrocatechol als idealer Modellligand, mit einem K_D -Wert im μM Bereich, wohingegen die Zuckermoleküle und der Zuckeralkohol K_D -Werte im mM Bereich aufwiesen. Der Grund für den stabileren Komplex mit 4-Nitrocatechol ($K_D = 4,7 \pm 0,05 \mu\text{M}$, Abbildung 12 A) könnte an der günstigen Geometrie der benachbarten Hydroxylgruppen als Substituenten an einem planaren, Kohlenstoff liegen, aber auch an dessen aromatischem Charakter und der elektronenziehenden sowie resonanzstabilisierenden Nitrogruppe (Pizer & Babcock, 1977). Bei der Titration von 36Bpa-NFW mit Sorbit konnte ein K_D -Wert von $2,1 \pm 0,1 \text{ mM}$ bestimmt werden (Abbildung 12 B). Als Polyalkohol weist Sorbit sechs Hydroxylgruppen auf, die frei drehbar sind und daher mehr reaktive *cis*-Diol-Gruppen bieten als die Pyranosen D-Galactose, D-Mannose und D-Glucose. Die zyklischen Halbacetale D-Galactose und D-Mannose besitzen abgesehen vom anomeren Kohlenstoff-Atom nur je ein *cis*-Diol-Gruppierung. Trotzdem unterscheidet sich der K_D -Wert von D-Galactose ($18 \pm 2,4 \text{ mM}$) und D-Mannose ($49,6 \pm 4,9 \text{ mM}$) ungefähr um den Faktor 2 (Abbildung 12 C, rot und grün). Dies weist auf eine Selektivität der molekularen Erkennung durch die Lipocalin-Bindetasche hin.

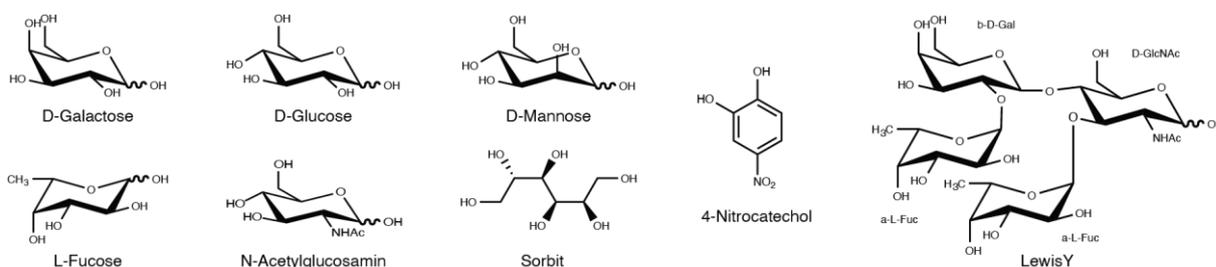


Abbildung 11: Strukturen der für die Titration verwendeten Zucker und Zuckerderivate.

Die Titration mit D-Glucose, deren Hydroxylgruppen alle in der *trans*-Konfiguration vorliegen, mit Ausnahme der anomeren Kohlenstoffposition und dessen Nachbarn, führte zum höchsten

K_D -Wert von 118 ± 23 mM. Somit weist 36Bpa-NFW eine um den Faktor 6,5 geringere Affinität zu D-Glucose auf als zu D-Galactose (Abbildung 12 C). Die Titration der Tyr-Variante (36Y-NFW) sowie von wtLcn2 mit D-Galactose ergab erwartungsgemäß keine nennenswerten Bindungssignale (Edwardraja *et al.*, 2017).

In weiterer Folge wurde die Affinität von 36Bpa-NFW zu dem Tetrasaccharid LewisY (α -Fuc-(1 \rightarrow 2)- β -Gal-(1 \rightarrow 4)-(α -Fuc-[1 \rightarrow 3])-GlcNAc) sowie zu L-Fucose und *N*-Acetylglucosamin (Abbildung 12) ermittelt. Obwohl die L-Fucose ähnlich wie die D-Galactose ein *cis*-Diol aufweist und LewisY als Tetrasaccharid sogar drei *cis*-Diole trägt, konnte kein Unterschied im Bindungsverhalten von 36Bpa-NFW gegenüber dem Kontrollprotein 36Y-NFW festgestellt werden (Abbildung 12 D und F). Auch bei *N*-Acetylglucosamin, einem Derivat der D-Glucose mit einem Acetylrest an C₂, wurde keine relevante Affinität von 36Bpa-NFW im Vergleich zu der Tyrosin-Variante detektiert (Abbildung 12 E).

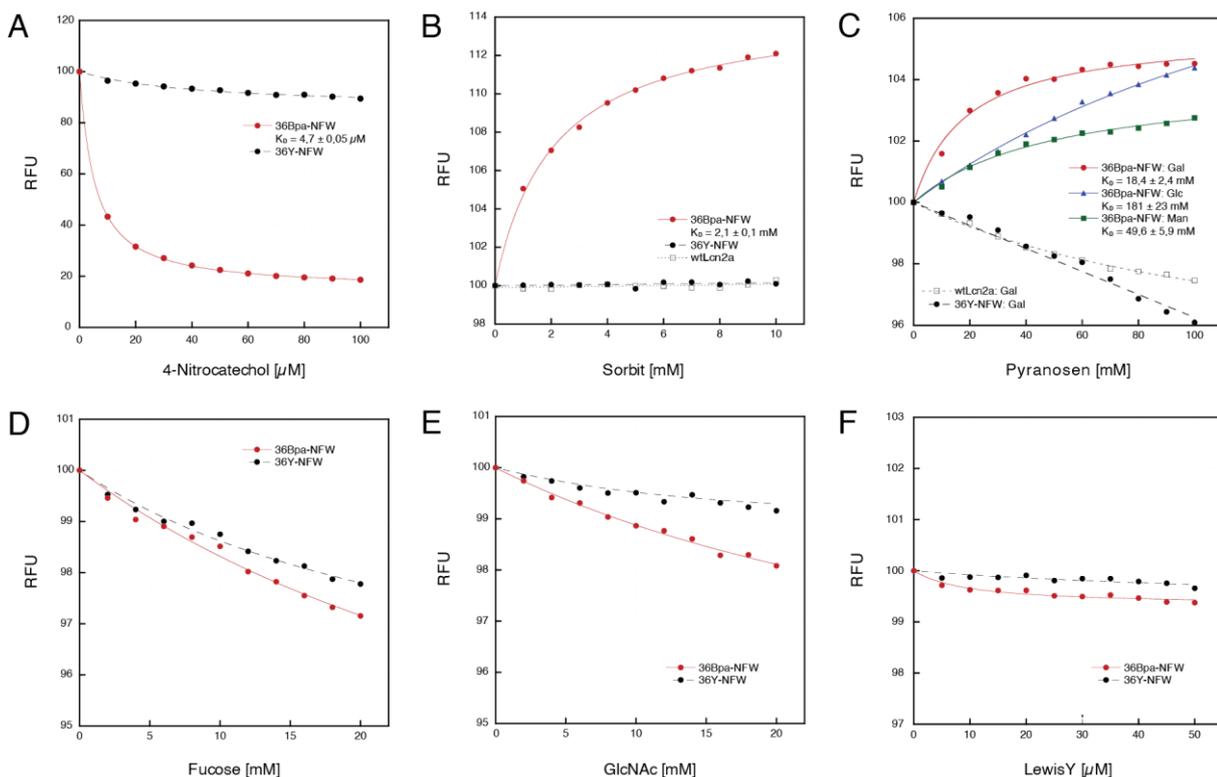


Abbildung 12: Fluoreszenztitration von 36Bpa-NFW mit Zuckern und Zuckerderivaten. Eine gereinigte Proteinlösung (1 µM) von 36Bpa-NFW oder 36Y-NFW in CHES / EDTA Puffer pH 8,5 wurde mit dem Liganden – als konzentrierte Stammlösung in dem selben Puffer – schrittweise titriert und die Veränderung der Proteinfluoreszenz bei 345 nm nach Anregung bei 294 nm in einen Spektrofluorimeter LS50B (Perkin Elmer) verfolgt. Zur Bestimmung des K_D -Wertes wurde die Anfangsfluoreszenz auf 100 % normiert, um den *Inneren-Filter-Effekt* korrigiert und die Messwerte wurden durch nicht-lineare Regression angepasst. (A) Titration von 36Bpa-NFW im Vergleich zu der Tyrosin-Variante mit 4-Nitrocatechol. (B) Titration von 36Bpa-NFW, 36Y-NFW und wtLcn2 mit Sorbit. (C) Titration von 36Bpa-NFW mit D-Galactose (Gal), D-Glucose (Glu) und D-Mannose (Man) sowie der Tyrosin-Variante und von wtLcn2 (als negativ Kontrolle) gegen D-Galactose. (D) Titration von 36Bpa-NFW und 36Y-NFW mit L-Fucose. (E) Titration von 36Bpa-NFW und der Tyrosin-Variante mit *N*-Acetylglucosamin (GlcNAc). (F) Titration von 36Bpa-NFW und 36Y-NFW mit LewisY.

Um weitere Beweise für die Bildung des kovalenten Komplexes zwischen 36Bpa-NFW und dem *cis*-Diol des Liganden zu sammeln, wurden ein Massenspektrum mittels ESI-MS unter nativen Bedingungen aufgenommen (Abschnitt 2.6.2 (Loney & Heck, 2017)).

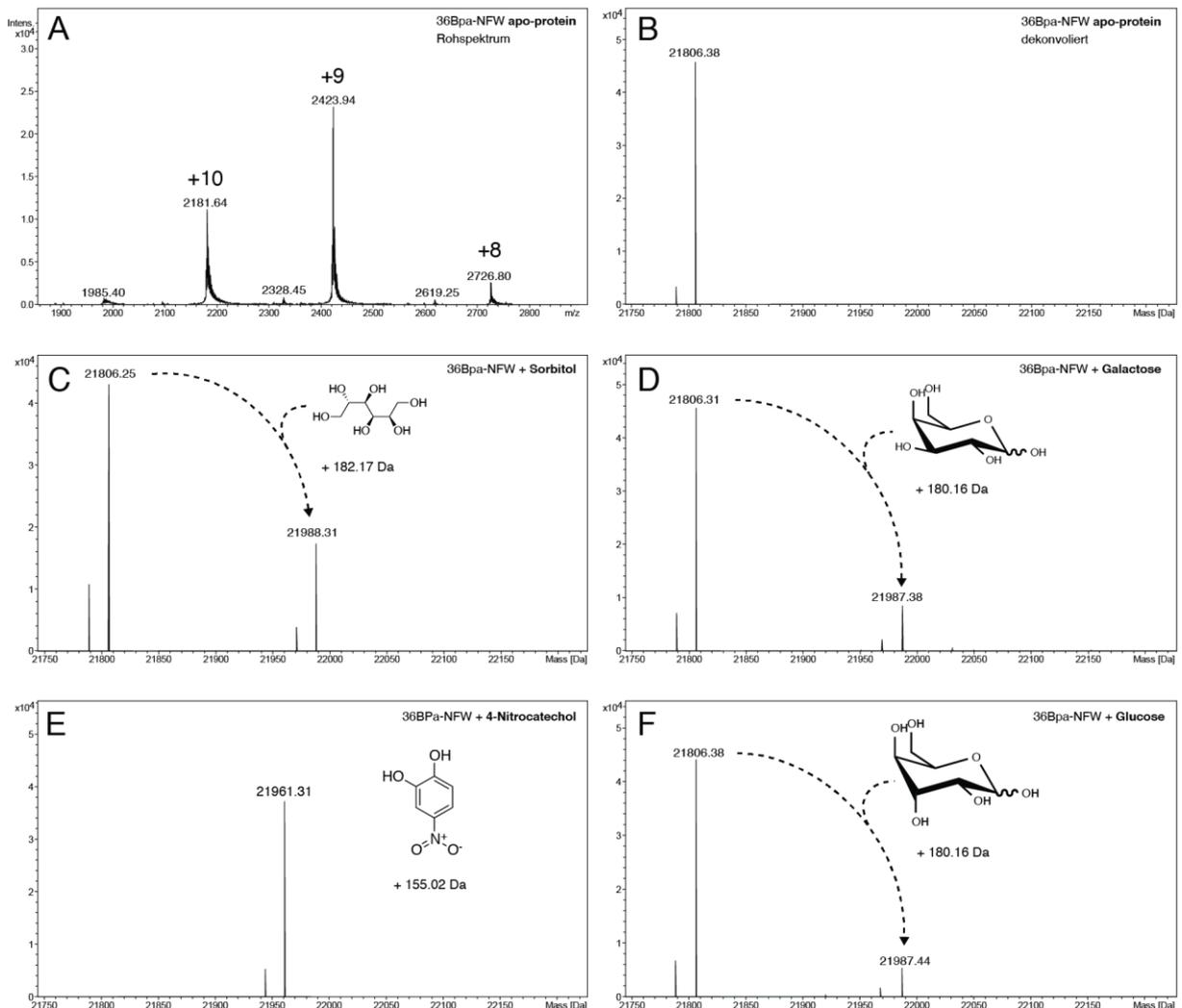


Abbildung 13: Native ESI-Massenspektren von 36Bpa-NFW und dessen kovalenten Komplexen mit Liganden. (A) Unter nativen ESI-Bedingungen tritt das Protein mit nur 8 – 10 Ladungszuständen bei einem m/z Verhältnis von ≈ 2500 , in die Gasphase ein. (B) Dekonvoliertes Massenspektrum des Apo-Proteins ($MW_{\text{calc}} = 21842,59 \text{ Da}$; $MW_{\text{obs.}} = 21806,38 \text{ Da}$; $\Delta m = -36,33 \text{ Da}$). Der Verlust von zwei Wassermolekülen weist auf intramolekulare Kondensationsreaktionen zwischen der Boronat-Seitenkette und benachbarten Alkohol- oder Amin-Seitenketten hin (Brustad *et al.*, 2008). (C-F) Natives ESI-MS von 36Bpa-NFW in Gegenwart von Sorbit (C), D-Galactose (D), 4-Nitrocatechol (E) und D-Glucose (F), jeweils in einem molaren Verhältnis von 1:1. In jedem Fall wurde ein Signal detektiert, das dem kovalenten Komplex aus Protein und Ligand entspricht, neben dem Signal für das Apo-Protein (mit Ausnahme von 4-Nitrocatechol).

Die Bezeichnung „nativ“ bezieht sich hierbei auf die Probenvorbereitung, bei der im Gegensatz zur klassischen Massenspektrometrie bei denen die Proteine mit organischen Lösungsmitteln und Säuren denaturiert werden, bleiben hier die Proteine intakt. Um den Protein-Ligand-Komplex zu detektieren, wurde vorab die Proteinlösung in einem Ammoniumbicarbonat Puffer, pH 8,85 im molaren Verhältnis von 1:1 mit dem Liganden (4-Nitrocatechol, Sorbit, D-Galactose, D-Glucose) inkubiert. Das Spektrum zeigte bei allen Liganden einen Signal-Peak

für den Komplex (Abbildung 13). In Gegenwart von 4-Nitrocatechol (Abbildung 13 E) wurde nur dieses prominente Signal detektiert, in Gegenwart von Sorbit, D-Galactose, D-Glucose (Abbildung 13 C, D und F) auch das Signal für das Apo-Protein detektiert. Die Signalverhältnisse stehen im Einklang mit den unterschiedlichen Affinitäten der Borocalin-Variante gegenüber dieser Liganden (Edwardraja *et al.*, 2017).

Im Anschluss an die Evaluierung der Bindungsaffinität gegenüber Zuckern und Zuckerderivaten der Borocalin-Variante Bpa-NFW, die durch rationales Engineering gewonnen wurde, erfolgten entsprechende Analysen der Boroacilin-Varianten 36Bpa-R2C2, 36Bpa-R2C2.1 und 36Bpa-R2C2.7, die aus den Selektionskampagnen gegen LewisY resultierten. Allerdings konnte bei keiner dieser drei Varianten mittels Fluoreszenztitration eine Affinität gegenüber dem Tetrasaccharid LewisY festgestellt werden (Abbildung 14 D).

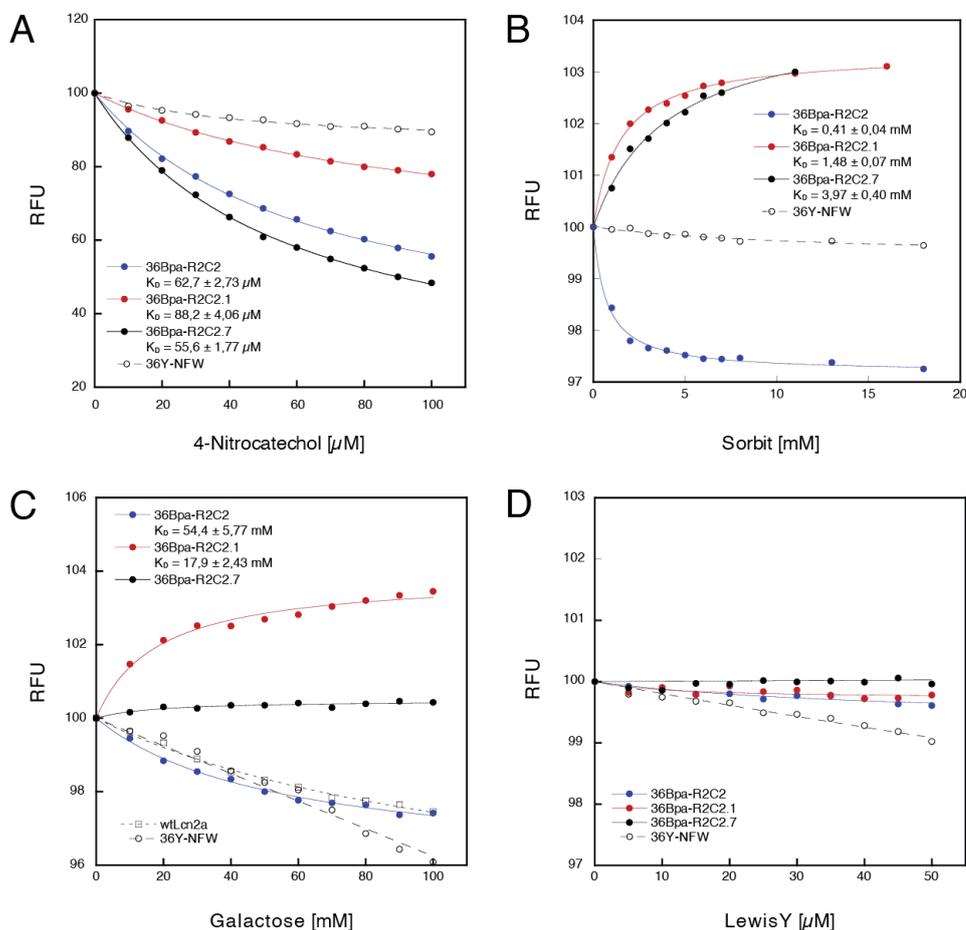


Abbildung 14: Fluoreszenztitration der aus der Selektion gegen LewisY resultierenden Borocalin-Varianten mit Zuckern und Zuckerderivaten. Eine gereinigte Proteinlösung (1 µM) von 36Bpa-R2C2, 36Bpa-R2C2.1, 36Bpa-R2C2.7 oder 36Y-NFW in CHES / EDTA Puffer pH 8,5 wurde mit dem Liganden – als konzentrierte Stammlösung in dem selben Puffer – schrittweise titriert und die Veränderung der Proteinfluoreszenz bei 345 nm nach Anregung bei 294 nm in einem Spektrofluorimeter verfolgt. Zur Bestimmung des K_D -Wertes wurde die Anfangsfluoreszenz auf 100 % normiert, um den *Innerer-Filter-Effekt* korrigiert und die Messwerte wurden durch nicht-lineare Regression angepasst. Titration der Varianten im Vergleich zu 36Y-NFW mit (A) 4-Nitrocatechol, (B) Sorbit, (C) D-Galactose (Gal) und (D) LewisY.

Zudem zeigte keine der Borocalin-Varianten eine höhere Affinität gegenüber dem Monosaccharid D-Galactose als 36Bpa-NFW. Die Variante 36Bpa-R2C2.1 ($K_D = 17,9 \pm 2,4$ mM) wies eine vergleichbare Affinität und 36Bpa-R2C2 ($K_D = 54,4 \pm 5,8$ mM) eine um den Faktor 3 niedrigere Affinität auf. Bei der Titration von 36Bpa-R2C2.7 mit D-Galactose konnte hingegen das charakteristische Sättigungsverhalten der Fluoreszenzsignals, wie es bei einer Protein-Liganden-Interaktion der Fall wäre, nicht beobachtet werden (Abbildung 14 C). Zusätzlich war bei diesen drei Borocalin-Varianten im Fall der Titration mit D-Galactose die relative Veränderung des Fluoreszenzsignals im Vergleich zu 36Bpa-NFW um die Hälfte geringer. Die Affinität gegenüber den Zuckeralkohol Sorbit erwies sich interessanterweise bei 36Bpa-R2C2 ($K_D = 0,41 \pm 0,04$ mM) am höchsten und war um den Faktor 5 beziehungsweise den Faktor 4 höher als bei 36Bpa-NFW und 36Bpa-R2C2.1 ($K_D = 1,38 \pm 0,07$ mM). Die niedrigste Affinität gegenüber Sorbit zeigte hingegen Bpa-R2C2.7, mit einem K_D -Wert von $3,97 \pm 0,40$ mM (Abbildung 14 B). Des weiteren zeigten die neuen Borocalin-Varianten aus den Selektionskampagnen die niedrigste Affinität gegenüber 4-Nitocatechol (Abbildung 14 A). Dies entsprach den Erwartungen, da sich 4-Nitocatechol und das für die Selektion verwendete LewisY strukturell erheblich unterscheiden (Abbildung 11). Angesichts dieser Befunde ist ein neuer experimenteller Ansatz erforderlich, um Borocalin-Varianten mit Affinität gegenüber LewisY zu selektieren.

3.1.3 Kristallstruktur der Borocalin-Variante Lcn2a36Bpa-NFW im Komplex mit 4-Nitocatechol

Zur strukturellen Analyse des Bindungsmechanismus des Borocalins und aufzuklären, ob die Bpa-Seitenkette als Boronsäure (trigonal planare Struktur) oder als Boronat (tetraedrische Struktur) mit dem *cis*-Diol einen Diester bildet, wurde Lcn2a36Bpa-NFW im Komplex mit Liganden kristallisiert (Abschnitt 2.7). Trotz der strukturellen Unterschiede zu Zuckermolekülen eignete sich das Modellsubstrat 4-Nitocatechol aufgrund des niedrigen K_D -Wertes am besten dafür. Die Borocalin-Variante wurde gentechnisch hergestellt und über das *Strep*-tag II mittels Affinitätschromatographie und anschließend durch Anionenaustauschchromatographie und Größenausschlusschromatographie gereinigt. Abschließend wurde die Proteinlösung gegen 20 mM CHES/NaOH, pH 8,5 dialysiert. Die Proteinlösung wurde zur Komplexbildung im molaren Verhältnis von 1:3 mit 4-Nitocatechol gemischt. Der Komplex wurde daraufhin an Dr. Andreas Eichinger, Lehrstuhl für Biologische Chemie, übergeben und für Kristallisationsexperimente im sitzenden Tropfen mit 5 x 96 bei unterschiedlichen Bedingungen (Vector Screen 1 bis 5) verwendet.

Die einzelnen Fällungsansätze wurden mit Hilfe eines Pipettierroboters je 400 nl Komplex und je 400 nl Reservoir pipettiert und bei 20 °C inkubiert. Nach mehreren Wochen wurde ein Kristall, der in Gegenwart von 2 M Ammoniumsulfat gewachsen war, entdeckt. Zur Optimierung hinsichtlich Größe und Beugungsfähigkeit wurde ein Fine Screen mit hängenden

Tropfen bei 20 °C angesetzt. Der vielversprechendste Kristall, gezüchtet mit 1,6 M Ammoniumsulfat, wurde unter Zugabe von 30 % (v/v) Glycerol eingefroren und an der Synchrotronstrahlenquelle 14.1 des BESSY in Berlin vermessen.

Der Röntgen-Beugungsdatensatz für diesen Kristall wies eine Auflösung von 1,98 Å mit der Raumgruppe C2 und zwei Protein Molekülen in der asymmetrischen Einheit auf. Die Prozessierung der Daten sowie das Lösen der Struktur durch molekularen Ersatz unter Verwendung von Apo-36Bpa-NFW (PDB ID: 5MHH; Edwardraja *et al.*, 2017) als Suchmodell wurde von Dr. Andreas Eichinger durchgeführt. Die Elektronendichte des Borocalins war von Ser⁵ bis Lys¹⁸⁸ inklusive Bpa an Position 36 mit dem kovalent gebunden 4-Nitrocatechol sichtbar. Des weiteren zeigte die Kristallstruktur, das für die Familie der Lipocaline charakteristische Faltungsmotiv, bestehend aus acht antiparallelen β -Faltblattstängen, die ein β -Fass bilden mit einer angelagerten α -Helix. Trotz der unterschiedlichen Raumgruppen und somit unterschiedlicher Kristallpackung weist der Borocalin Ligand-Komplex hohe Deckungsgleichheit zum Suchmodell auf. Bei der Überlagerung der strukturell konservierten 58 C α -Atome (Skerra, 2000) betrug die Abweichung (RMSD) bloß 0,225 Å. Eine nur leicht höhere Abweichung (RMSD = 0,445 Å) zeigte sich bei der Überlagerung von allen 174 äquivalenten C α -Atomen, abgesehen vom *Strep*-tag II, auf. Dies bedeutet, dass zusätzlich zum β -Fass auch die Konformationen der vier strukturell variablen Loops, die die Bindetasche formen (Skerra, 2000), konserviert sind. Nur die Konformation der Aminosäurenreste des *Strep*-tag II Peptides variierten im Vergleich zum Suchmodell um 4,1 Å an der C α -Position. Interessanterweise ragte das ein *Strep*-tag II eines benachbarten Moleküls in der Kristallpackung des Apo-Borocalins (Edwardraja *et al.*, 2017) in die leere Bindetasche der Apo-Struktur, was im Fall der Komplexbildung (Sommer *et al.*, 2020) nicht möglich ist (Abbildung 15 D).

In dem nun kristallisierten Komplex ist die artifizielle Bpa-Seitenkette deutlich zu sehen. Sie bildet einen zyklischen Komplex mit der 1,2-*cis*-Diol Gruppe des 4-Nitrocatechols (Abbildung 15 B). Im Gegensatz zur Kristallstruktur des Apo-Borocalins, bei dem die Bpa-Seitenkette mit trigonal-planarer Boronsäuregruppe vorliegt (Abbildung 15 A), erscheint in der Kristallstruktur des Protein/Ligand-Komplexes die Bpa-Seitenkette in der tetragonalen Struktur (Abbildung 15 B). Die freie Hydroxylgruppe bildet — mit einem Abstand von 2,7 Å bei nicht idealem Wasserstoffbrücken-Bindungswinkel — bildet eine polare Wechselwirkung mit dem Asn¹³⁴. Es scheint, als würde diese Interaktion die tetraedrische Konfiguration, die die Bpa-Seitenkette eher bei alkalischem pH-Wert einnimmt, stabilisieren. Im Vergleich zur Apo-Kristallstruktur verursacht diese polare Interaktion eine leichte Verschiebung der Bpa-Seitenkette in Richtung Asn¹³⁴ (Abbildung 15 C). Zudem wird die Orientierung von 4-Nitrocatechol über zwei Wasserstoffbrückenbindungen zwischen der Seitenkette Ser⁶⁸ (2,2 Å Distanz) und der von Arg⁸¹ (2,9 Å Distanz) und der Nitro-Gruppe stabilisiert.

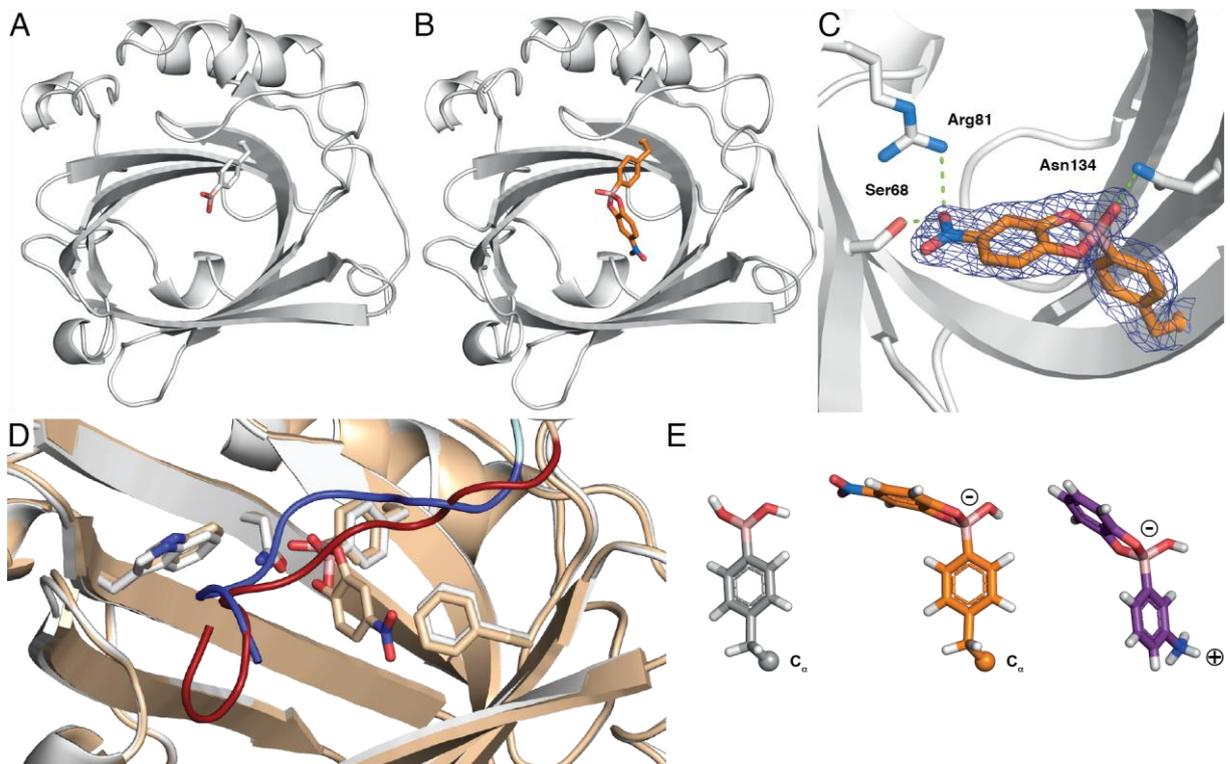


Abbildung 15: Kristallstrukturen der Borocalin-Variante Lcn2a36Bpa-NFW. (A) Apo-Borocalin in 10 mM MOPS/NaOH pH 7,4, kristallisiert im hängenden Tropfen mit 2,5 M Ammoniumsulfat als Fällungsreagenz bei 20 °C (PDB ID: 5MHH; Edwardraja *et al.*, 2017). (B) Der hier gelöste Borocalin-Nitrocatechol Komplex, kristallisiert im hängenden Tropfen mit 1,6 M Ammoniumsulfat als Fällungsreagenz bei 20 °C. (C) Nahansicht der Bindetasche des Borocalin/Nitrocatechol Komplexes (PDB ID: 6QMU). Zu sehen ist der zyklische Diester zwischen der Boronsäure in ihrer tetraedrischen Konfiguration und 4-Nitrocatechol mit eingeleiteter Elektronendichte. Drei Wasserstoffbrücken und/oder polare Wechselwirkungen sind als grüne unterbrochenen Linie dargestellt. (D) Überlagerung des Apo-Borocalins (grau) und des Borocalin/Nitrocatechol Komplexes (beige). Zu sehen ist das *Strep*-tag II in der Bindetasche eines benachbarten Moleküls, wobei dessen Lage deutlich variiert (blau für die Apo-Kristallstruktur, rot für die Kristallstruktur des Komplexes). (E) Struktureller Vergleich von ungebundenem Bpa (grau), dem Diester aus Bpa, 4-Nitrocatechol (orange) und 2-(3-aminophenyl)benzo-1,3,2-dioxaborol (violett) (Vega *et al.*, 2010). Das C α -Atome ist jeweils als Kugel gekennzeichnet.

Die hier gelöste Kristallstruktur zeigt zum ersten Mal den zyklische Diester zwischen einer solvatisierten Boronsäure in der Bindetasche eines Proteins mit einem *cis*-Diol in wässriger Umgebung. Denn obwohl seit über einem Jahrhundert bekannt ist, dass Boronsäuren oder ihre Derivate Ester mit Zuckern beziehungsweise mit *cis*-Diolen bilden können (v. Vargha, 1933), gibt es wenig strukturelle Informationen darüber und diese basieren hauptsächlich auf Konduktivitäts-, spektroskopischen (Böeseken, 1949; Springsteen & Wang, 2002) und NMR-Messungen (van den Berg *et al.*, 1994). Eine publizierte Kristallstruktur zeigt einen niedermolekularen Komplex aus *m*-Aminophenylboronat und Catechol 2-(3-aminophenyl)benzo-1,3,2-dioxaborol. Auch hier ist das Boronat Teil eines Fünfrings, der mit Benzol anelliert ist (Abbildung 15 E), was zu einem niedrigeren pK_s der Boronsäure führt. Hier wird jedoch das Hydroxylgruppe des tetraedrischen Boronats durch die positiv geladene Anilium-Gruppe als inneres Salz stabilisiert (Vega *et al.*, 2010). Weitere Kristallstrukturen sind von Komplexen aus Enzymen und Boronsäure-Inhibitoren bekannt, bei denen mit Glycerol ein Ester gebildet wird (Lei *et al.*, 2016). Dies geschah offenbar nicht aufgrund einer spezifischen Affinität zum

Liganden, sondern aufgrund der Vorbereitung für die Röntgenstrukturanalyse. Nach der erfolgreichen Ko-Kristallisation von 36Bpa-NFW und 4-Nitrocatechol wurde versucht, 36Bpa-NFW im Komplex mit L-Ribose, D-Galactose und LewisY zu kristallisieren. Dafür wurde die konzentrierte Proteinlösung zur Komplexbildung im molaren Verhältnis von 1:20 mit L-Ribose, 1:33 mit D-Galactose oder 1:3 mit LewisY gemischt. Die Kristallisationsversuche erfolgten im sitzenden Tropfen mit 5 x 96 unterschiedlichen Bedingungen (Vector Screen 1 bis 5). Die Fällungsansätze wurden wie zuvor mit Hilfe eines Pipettierroboters — für den Kristallisationsversuch mit L-Ribose und D-Galactose je 600 nl Proteinlösung und 600 nl Reservoirlösung und für den Kristallisationsversuch mit LewisY je 400 nl Proteinlösung und 400 nl Reservoirlösung – pipettiert und bei 20 °C inkubiert. Die dabei erhaltenen Kristalle erwiesen sich allerdings als Zucker- oder Salzkristalle.

3.1.4 Rationales Engineering ausgehend von Lcn2a36Bpa-NFW

Da es scheint, als würde die polare Wechselwirkung der Asn¹³⁴-Seitenkette mit der Bpa-Seitenkette deren tetraedrische Konfiguration stabilisieren, wurde die Rolle der Asn¹³⁴-Seitenkette hinsichtlich deren Einfluss auf die Affinität des Borocalins gegenüber Nitrocatechol und D-Galactose evaluiert. Hierfür wurde Asn¹³⁴ zu Ala mutiert — in Anlehnung an das *Alanin-Scanning* (Morrison & Weiss, 2001). Die Ala-Seitenkette ist eine inerte Methylgruppe, die keine Interaktion mit Bpa eingehen kann und erfahrungsgemäß die Sekundärstruktur des Proteins nicht beeinflusst. Diese Variante wurde 36Bpa-AFW genannt. Die Borocalin-Varianten, 36Bpa-AFW und 36Bpa-NFW wurden wie bereits beschrieben gentechnisch produziert (Abschnitt 2.4.1.1), unter Nutzung eines C-terminalen His₆-Tags wurden beide mittels Immobilisierter-Metallionen-Affinitätschromatographie (IMAC) (Abschnitt 2.5.2) und AEX (Abschnitt 2.5.3) gereinigt. Aus der Produktion der Varianten im 2-L-Kulturmaßstab resultierten nach der Reinigung 4,7 mg 36Bpa-NFW und 1,8 mg 36Bpa-AFW. Die gereinigten Proteine wurden im Anschluss für die Ermittlung der Affinität mittels Fluoreszenztitration verwendet. Hierbei zeigte 36Bpa-AFW ($K_D = 29,3 \pm 0,72 \mu\text{M}$) eine 6-fach geringere Affinität zu 4-Nitrocatechol als 36Bpa-NFW ($K_D = 4,72 \pm 0,05 \mu\text{M}$), das auf die fehlende Amidgruppe zur Stabilisierung des Boronat-Ions hinwies (Abbildung 17 A). Bei der Titration mit D-Galactose hingegen konnte kein Unterschied des K_D -Wertes zwischen beiden Borocalin-Varianten festgestellt werden (Abbildung 17 B). Beide zeigten eine charakteristische Sättigungskurve im Gegensatz zu 36Y-NFW (Abbildung 17 A und B). Das ließ darauf schließen, dass nicht nur die Boronsäure mit ihrer Reaktivität in der tetraedrischen Konfiguration einen Einfluss auf die Affinität hat, sondern auch die Bindetasche des Lipocalins bei der Erkennung des Liganden eine wichtige Rolle spielt.

Versuche zur strukturellen Aufklärung des Borocalin/Galactose-Komplexes durch Kristallisation waren im Gegensatz zum Borocalin/Nitrocatechol-Komplex nicht erfolgreich. Daher konnte ein rationales Engineering zur Steigerung der Affinität nur bedingt durchgeführt

werden. Die relative Fluoreszenzänderung der Proteinlösung bei der Titration mit D-Galactose betrug nur 6 %. Daraus resultierte die Frage, ob eine verbesserte Affinität wegen der geringen relativen Fluoreszenzänderung überhaupt detektiert werden konnte.

Zur Steigerung der relativen Fluoreszenzänderung aufgrund der Komplexbildung mit D-Galactose wurden zwei aromatische Aminosäuren (Phe⁵² und Tyr¹³²) in der Bindetasche des Borocalins ausgewählt, die vermutlich mit dem Liganden interagieren, und zu Trp mutiert. Trp reagiert besonders sensibel auf Veränderungen der direkten Umgebung (Lakowicz, 2013). Für die Borocalin-Variante 36Bpa-NW-F52W wurde die Aminosäure Phe⁵², welche sich in direkter Umgebung zu Bpa befindet, zu Trp mutiert (Abbildung 16 B) und für die Variante 36Bpa-NFW-W79Y wurde Trp⁷⁹ zu Tyr (Abbildung 16 C) mutiert. Im Gegensatz dazu wurde auch eine Variante generiert (36Bpa-NFW-Y132W) welche die Bindung von größeren Molekülen als 4-Nitrocatechol erlauben sollte. Dafür wurde Tyr¹³² zu Trp (Abbildung 16 D) mutiert.

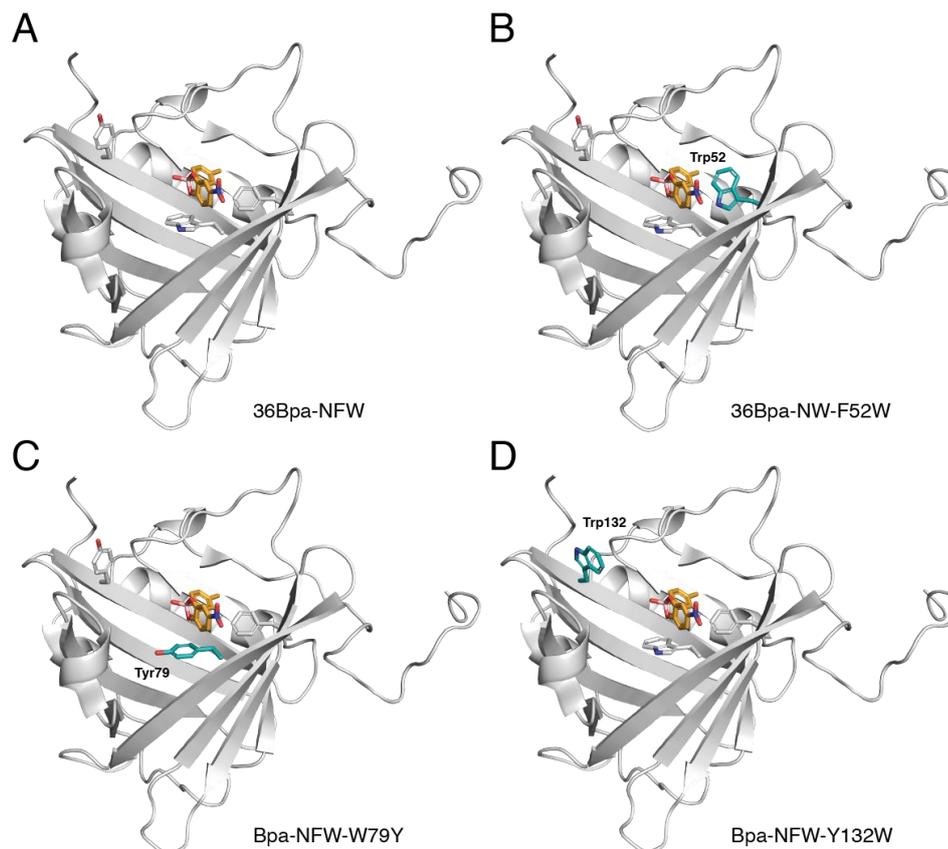


Abbildung 16: Als Cartoon dargestellte Modelle der Borocalin-Varianten im Komplex mit 4-Nitrocatechol. (A) 36Bpa-NFW, als Stabmodell sind die zur Mutagenese ausgewählten Aminosäurepositionen (grau) und der Boronsäure-Nitrocatechol Komplex (orange) dargestellt. Die mutierten Aminosäurepositionen der Borocalin-Varianten 36Bpa-NW-F52W (B), 36Bpa-NFW-W79Y (C) und 36Bpa-NFW-Y132W sind in Türkis gekennzeichnet.

Aus der gentechnischen Produktion im 2-L-Kulturmaßstab und Reinigung der drei Borocalin-Varianten resultierten 5,6 mg 36Bpa-NW-F52W sowie 0,7 mg 36Bpa-NFW-W79Y und 3,6 mg 36Bpa-NFW-Y123W.

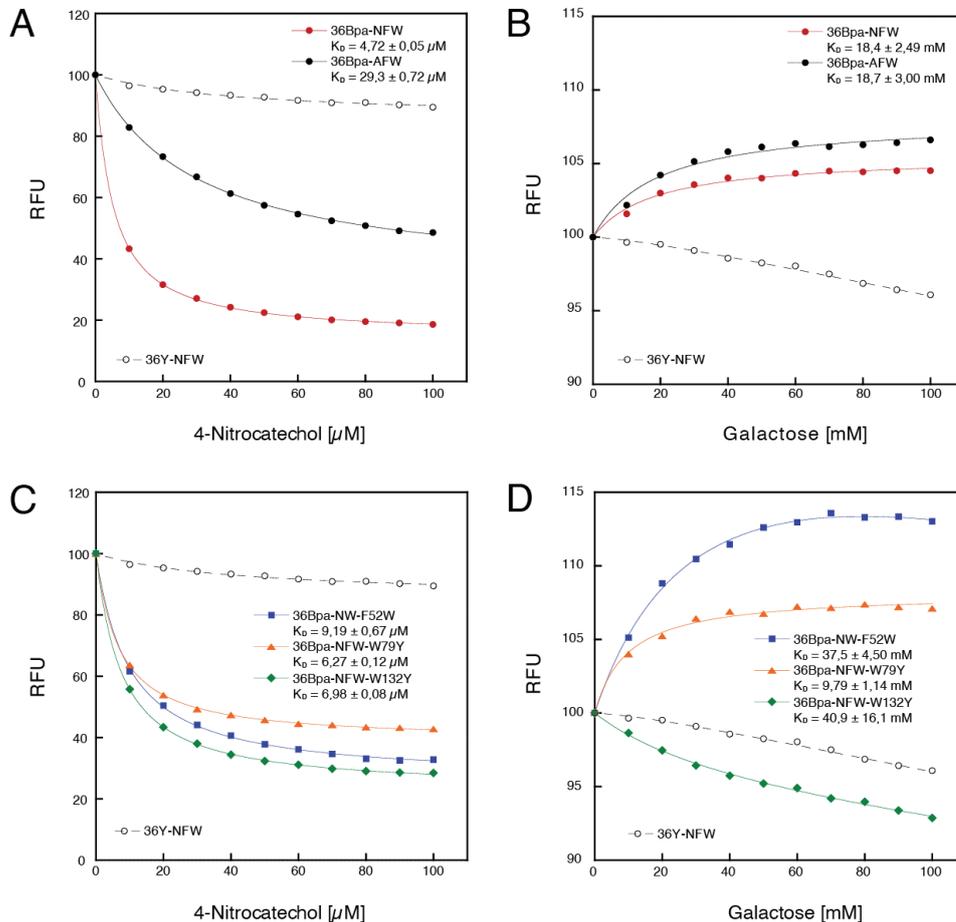


Abbildung 17: Fluoreszenztitration der durch rationales Engineering erzeugten Borocalin-Varianten mit Diolen. Eine gereinigte Proteinlösung (1 μM) mit 36Bpa-NFW, 36Bpa-AFW, 36Bpa-NW-F52W, 36Bpa-NFW-W79Y oder 36Bpa-NFW-Y132W in CHES / EDTA Puffer pH 8,5 wurde mit dem Liganden – als konzentrierte Stammlösung in demselben Puffer – schrittweise titriert und die Veränderung der Proteinfluoreszenz bei 345 nm nach Anregung bei 294 nm in einem Spektrofluorimeter verfolgt. Zur Bestimmung des K_D -Wertes wurde die Anfangsfluoreszenz auf 100 % normiert, dann um den *Inneren-Filter-Effekt* korrigiert und die Messwerte durch nicht-lineare Regression angepasst. Titration der Varianten im Vergleich mit 36Y-NFW gegen (A und C) 4-Nitrocatechol oder (B und D) D-Galactose (Gal).

Die relative Fluoreszenzänderung bei der Komplexbildung zwischen 36Bpa-NW-F52W und D-Galactose konnte erwartungsgemäß verbessert werden (Abbildung 17 B und D), allerdings nahm die Aktivität ($K_D = 37,5 \pm 4,50 \mu\text{M}$) um den Faktor 2 im Vergleich zu der Ausgangsvariante, 36Bpa-NFW, ab (Abbildung 17 D). Die Titration mit 4-Nitrocatechol zeigte sowohl eine Abnahme der Affinität, ebenfalls um den Faktor 2, als auch eine Abnahme der relativen Fluoreszenzveränderung. Das wies auf eine sterische Beeinträchtigung bei der Komplexbildung durch die im Vergleich zu Phe deutlich größere Seitenkette von Trp hin (Abbildung 17 C). Bei der Borocalin-Variante 36Bpa-NFW-W79Y wurde beobachtet, dass sich die Affinität zu D-Galactose um den Faktor 2 erhöhte, die relative Fluoreszenzveränderung

jedoch sank. Somit wirkte sich die Mutation zwar vermutlich positiv auf die Komplexbildung aus, aber die Mutation von Trp⁷⁹ zu Tyr hatte eine Abnahme der relativen Fluoreszenzveränderung zur Folge. Auch eine Anregung bei 280 nm anstelle von 294 nm, der Wellenlänge, bei der neben der Fluoreszenz von Trp auch die von Tyr angeregt wurde, änderte an diesem Befund nichts. Bei der Titration dieser Variante mit 4-Nitrocatechol wurde ein K_D -Wert von $6,27 \pm 0,12 \mu\text{M}$ ermittelt, der nur gering über dem K_D -Wert der Ausgangsvariante 36Bpa-NFW lag (Abbildung 17 C). Ähnlich verhielt sich die Borocalin-Variante 36Bpa-NFW-Y132W bei der Titration mit 4-Nitrocatechol, was einen K_D -Wert von $6,98 \pm 0,08 \mu\text{M}$ ergab (Abbildung 17 C). Allerdings konnte in diesem Fall bei der Titration mit D-Galactose keine Sättigungsverhalten beobachtet werden (Abbildung 17 D). Das wies darauf hin, dass die ursprüngliche Aminosäure, Tyr¹³², zur Bindung von D-Galactose beitrug, aber keine Auswirkung auf die Erkennung von 4-Nitrocatechol hatte.

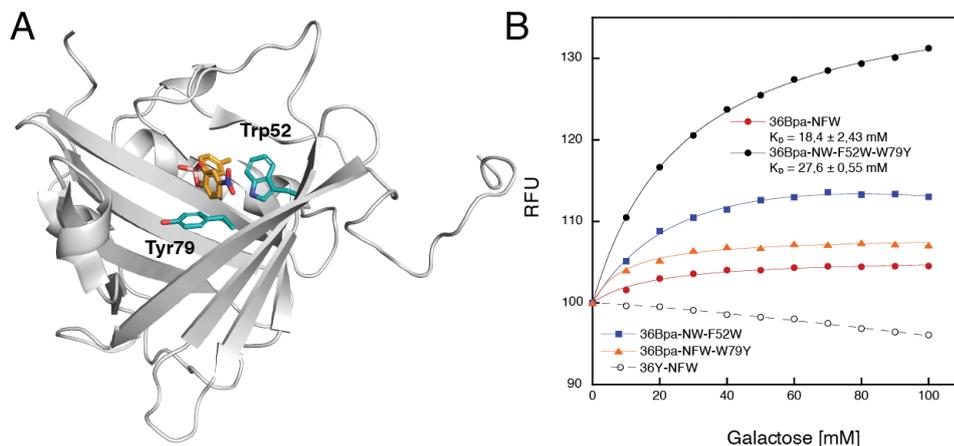


Abbildung 18: Modell der Borocalin-Doppelmutante und Fluoreszenztitration mit Galactose im Vergleich zu den Ausgangsvarianten. (A) Im Cartoon Modell ist die Borocalin-Variante (36Bpa-NW-F52W/W79Y) im Komplex mit 4-Nitrocatechol dargestellt. Im Stabmodell sind die beiden mutierten Aminosäuren (türkis) und der Boronsäure/Nitrocatechol Komplex hervorgehoben (orange). (B) Die gereinigte Proteinlösung ($1 \mu\text{M}$) in CHES / EDTA Puffer pH 8,5 wurde mit dem Liganden – als konzentrierte Stammlösung im selben Puffer – schrittweise titriert und die Veränderung der Proteinfluoreszenz bei 345 nm nach Anregung bei 294 nm in einem Spektrofluorimeter verfolgt. Zur Bestimmung des K_D -Wertes wurde die Anfangsfluoreszenz auf 100 % normiert, dann um den *Inneren-Filter-Effekt* korrigiert und die Messwerte durch nicht-lineare Regression angepasst. Titration der Doppelmutante und der Ausgangsvarianten im Vergleich zu 36Y-NFW mit D-Galactose.

Mit dem Ziel die gesteigerte relative Fluoreszenzänderung sowie die gesteigerte Affinität der Borocalin-Varianten 36Bpa-NW-F52W und 36Bpa-NFW-W79Y bei der Titration mit D-Galactose, zu vereinen, wurde die Doppelmutante 36Bpa-NW-F52W/W79Y (Abbildung 18 A) gentechnisch hergestellt und gereinigt. Die Proteinausbeute betrug hierbei 0,7 mg aus einer 2-L-Kultur. Durch Fluoreszenztitration mit D-Galactose wurde für diese Doppelmutante ein K_D -Wert von $27,6 \pm 0,55 \text{ mM}$ bestimmt. Die Affinität zu D-Galactose war damit höher als die von 36Bpa-NW-F52W (Abbildung 18 B). Die relative Fluoreszenzänderung konnte erwartungsgemäß gesteigert werden Da also die Affinität zu D-Galactose wie auch zu LewisY weder durch gerichtete Evolution noch durch rationales Protein-Engineering gesteigert werden

konnte, besteht die Hoffnung, dass eine reaktivere Bor-Aminosäure der Schlüssel zum Erfolg sein könnte.

3.1.5 Eine reaktivere Bor-Aminosäure für die selektive Bindung von Zuckermolekülen

Aus der Kristallstruktur von 36Bpa-NFW im Komplex mit 4-Nitrocatechol ist bekannt, dass Bpa als Boronat (Tetraeder) reaktiv ist. Bei neutralem pH-Wert liegt das chemische Gleichgewicht von Bpa (pKs ~9) jedoch auf der Seite der Boronsäure (Pizer & Babcock, 1977) und wird erst bei alkalischem pH auf die Seite des Boronats verschoben (Abbildung 19 A). Um die Reaktivität der Boronsäure zu steigern, wurde von Dr. Markus Anneser am Lehrstuhl für Biologische Chemie eine neuartige Bor-Aminosäure synthetisiert, die auch bei physiologischen pH-Wert als Boronat vorliegt (Abbildung 19 B). Diese neuartige Bor-Aminosäure (Borono-2,6-difluorphenyl-alanin, Bfa) unterscheidet sich von Bpa nur in zwei am Phenylring substituierten Fluoratomen. Die elektronenziehenden Eigenschaften des Fluors erniedrigen den pKs der Boronsäure, deren Gleichgewicht damit schon im neutralen pH-Wert auf die Seite der Boronats verschoben wird. Dabei sind die Fluoratome als Substituenten nur unwesentlich größer als Wasserstoffatome und sollten daher keinen wesentlichen sterischen Einfluss auf die Erkennung von Bfa durch die BpaRS haben.

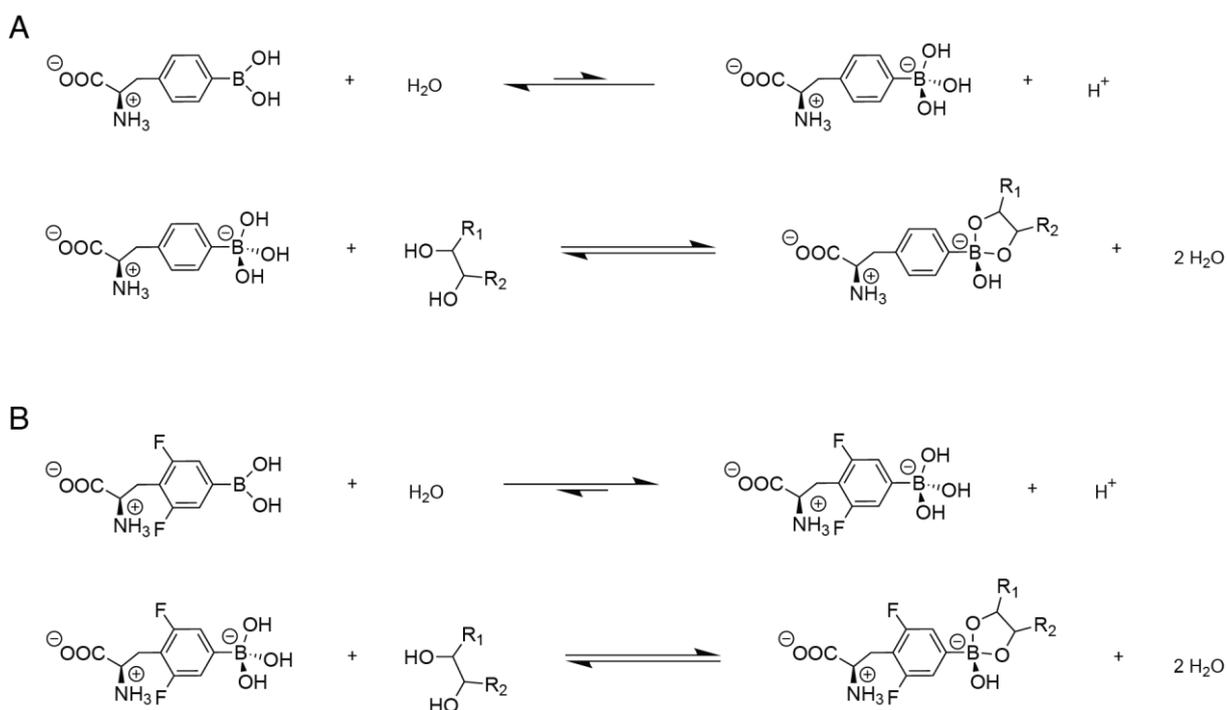


Abbildung 19: Chemisches Gleichgewicht der nicht-natürlichen Aminosäuren Bpa (A) und Bfa (B) in wässriger Lösung bei neutralem pH-Wert sowie bei der Komplexbildung mit einem *cis*-Diol.

Bei Bfa liegt aber der größere Anteil der Boronatgruppe als Tetraeder vor. Diese Konfiguration wird von der BpaRS nicht gebunden, denn für die selektive Bindung der Bpa-Seitenkette im aktiven Zentrum der BpaRS sind vier Wasserstoffbrückenbindungen verantwortlich (Schiefner & Skerra, 2015). Glu¹⁶² ist in *para*-Position, in einer Ebene zum Boronsäurerest in der trigonal planaren Struktur, ausgerichtet und bildet zwei Wasserstoffbrückenbindungen. Zusätzlich geht vom Ser³² und Gln¹⁰⁹, vermittelt über Wassermoleküle jeweils eine Wasserstoffbrückenbindung mit dem Boronsäurerest aus (Abbildung 20). Durch die Veränderung der Bpa-Seitenkette von der trigonalen- zur tetraedrischen Struktur ist die Ausbildung dieser Wasserstoffbrückenbindungen nicht mehr möglich. Hinzu kommt die elektrostatische Abstoßung der Boronatgruppe und Glu¹⁶². Jedoch liegt auch bei neutralem pH ein geringer Anteil von Bfa in trigonal planarer Struktur vor, die von der BpaRS erkannt wird.

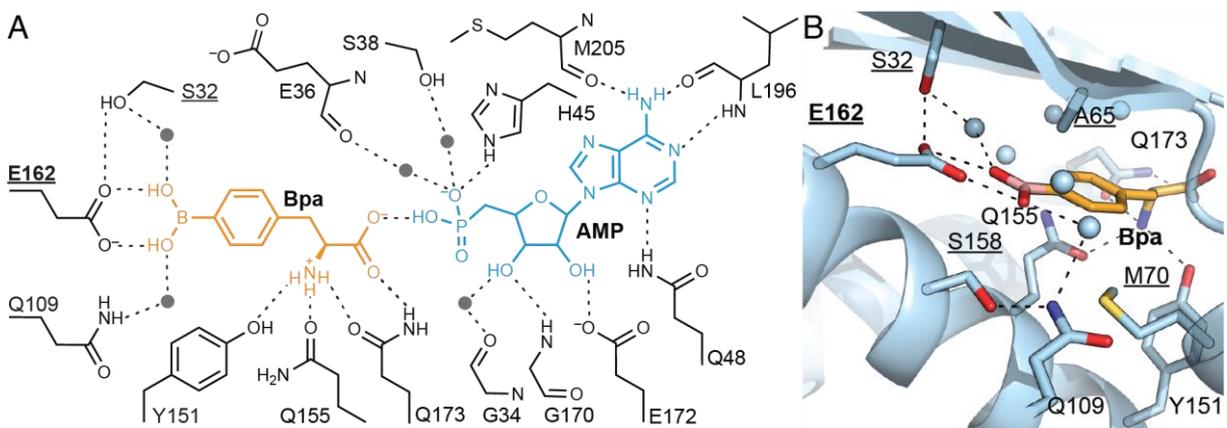


Abbildung 20: Erkennung von Bpa im aktiven Zentrum der BpaRS (Schiefner & Skerra, 2015), (A) schematisch und (B) als Kristallstruktur (PDB ID: 5N5U) dargestellt. Die in die Substraterkennung involvierten Aminosäuren sind unterstrichen.

Das bestätigte der Einbauversuch von Bfa mit Hilfe der BpaRS in das Reporterprotein eGFP(39a). Nur die erfolgreiche Suppression des Amber-Stop-Codons an Position 39 resultiert in einem eGFP-Protein mit voller Länge, dessen Fluoreszenz mit einem Durchflusszytometer (FACS-Gerät) detektiert werden kann (Kuhn *et al.*, 2010). Für dieses Experiment wurde das Plasmid pSB16.6c2 in den *E. coli* Stamm MG1655 transformiert und bei 37 °C in M9-Minimalmedium kultiviert (Abschnitt 2.4.1.2). Im Gegensatz zu Vollmedien, bei denen Trypton und Hefeextrakt verwendet werden, enthält das M9-Medium nur Salze und Glucose. Die für die Zelle benötigten Aminosäuren müssen endogen synthetisiert werden. Dadurch war eine gezielte Supplementierung der gewünschten (nicht-)natürlichen Aminosäuren möglich.

Erst nach 16-stündiger Gen-Expression von eGFP(39a) konnte ein Unterschied der Fluoreszenzintensität zwischen der Kultur, die in Gegenwart von Bfa kultiviert wurde, und der Kultur, die ohne zusätzliche Aminosäure-Supplementierung kultiviert wurde, detektiert werden

(Abbildung 21, türkis). Dies steht im Einklang mit der Hypothese, dass das chemische Gleichgewicht deutlich auf der Seite des Boronats liegt und nur ein geringer Anteil der nicht-natürlichen Aminosäuren als Boronsäure vorliegt. Um ein nicht-natürliches Protein mit Bfa zu produzieren, muss die Expressionszeit also erheblich verlängert werden. Zusätzlich wurde beobachtet, dass die von eGFP(39a), exprimiert in Gegenwart von Bfa, deutlich geringer war als die gemessene Fluoreszenzintensität von eGFP(39a), welches in Gegenwart von Tyr exprimiert wurde (Abbildung 21, türkis und grau). Die positiv Kontrolle, bei der eGFP(39a) in Gegenwart von Bpa exprimiert wurde zeigte die bestmögliche Fluoreszenzintensität unter Verwendung der BpaRS. Somit ist die BpaRS nicht zum effizienten Einbau von Bfa in das Lipocalin geeignet und bevor dieses mit Bfa funktionalisiert werden kann, muss eine für Bfa spezifische Aminoacyl-tRNA Synthetase konstruiert werden.

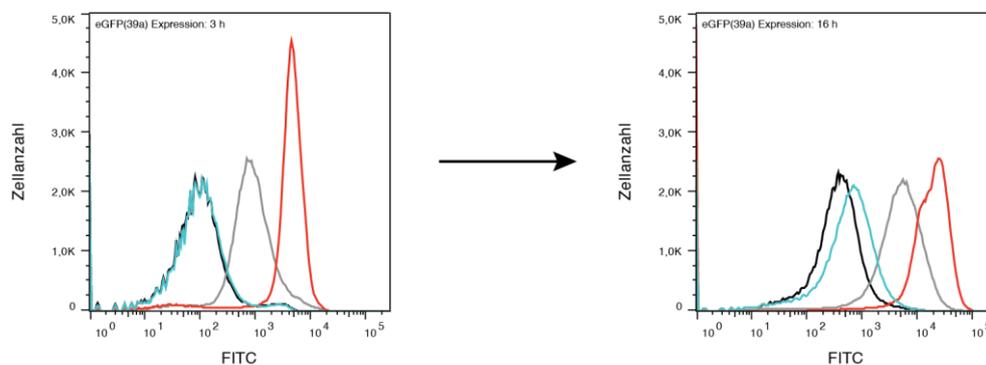


Abbildung 21: Evaluierung der BpaRS bezüglich der Erkennung von Bfa. Fluoreszenzsignal von eGFP(39a), erzeugt durch Suppression des Amber-Stop Codons in Gegenwart von 1 mM Tyr (grau), 1 mM Bpa (rot) oder 1 mM Bfa (cyan) mit Hilfe der BpaRS nach 3-stündiger beziehungsweise 16-stündiger Expression des Reporterproteins. In schwarz ist die Fluoreszenzintensität der Kultur ohne Supplementierung durch eine (nicht-)natürliche Aminosäure dargestellt. Für die Messung wurde die Spannung der Laser am FACS-Gerät wie folgt eingestellt: FITC = 550 V; FSC = 300 V; SSC = 500 V

3.2 Engineering und Selektion einer auf der Tyrosyl-tRNA Synthetase aus *Methanocaldococcus jannaschii* basierenden Aminoacyl-tRNA Synthetase für den Einbau von 7-Hydroxycoumarinalanin (Hca)

Für den Einbau einer neuartigen nicht-natürlichen Aminosäure in die Polypeptidkette und der damit verbundenen Funktionalisierung eines Proteins wird eine für die nnAA spezifische Aminoacyl-tRNA Synthetase benötigt. Dafür wird das aktive Zentrum einer Aminoacyl-tRNA Synthetase aus einem nicht verwandten Organismus mittels Zufallsmutagenese modifiziert und dann auf die molekulare Erkennung der nicht-natürlichen Aminosäuren über den daraus folgenden Einbau in ein Reporterprotein (eGFP(39a)) selektiert (Abschnitt 1.3) (Kuhn *et al.*, 2010). Ausgehend von der Tyrosyl-tRNA Synthetase aus *Methanocaldococcus jannaschii* (MjTyrRS) wurde versucht, eine Aminoacyl-tRNA Synthetase mit Spezifität für die nicht-natürliche Aminosäuren 7-Hydroxycoumarylanin (Hca) zu generieren. Hca ist eine stark fluoreszierende nicht-natürliche Aminosäure (Brun *et al.*, 2004) mit einer sterisch

anspruchsvollen Seitenkette an C β , bestehend aus einem bizyklischen Heteroaromaten mit einer OH-Gruppe an Position 7 (Abbildung 29).

3.2.1 Die SlonoMax Genbibliothek

Basierend auf der Kristallstruktur der Tyrosyl-tRNA Synthetase aus *Methanocaldococcus jannaschii* (*Mj*TyrRS, PDB ID: 1J1U) wurde von Prof. Dr. Arne Skerra und Sebastian Kuhn am Lehrstuhl für Biologische Chemie eine Genbibliothek für die molekulare Erkennung der nicht-natürlichen Aminosäure Hca geplant und von Sloning BioTechnology (jetzt MorphosSys AG, Puchheim) synthetisiert. Durch deren Technologie lassen sich gewünschte Mutationen an zuvor definierten Positionen in kontrollierter Häufigkeit und Verteilung generieren (Van den Brulle *et al.*, 2008). Diese Bibliothek umfasste insgesamt 17 Positionen, die vollständig oder nur teilweise randomisiert waren. Daraus ergab sich eine theoretische Komplexität von $1,3 \cdot 10^{22}$ Aminoacyl-tRNA Synthetase-Varianten. Die mutierten Positionen umfassten die erste, zweite und dritte Schale um das aktive Zentrum (Abbildung 22).

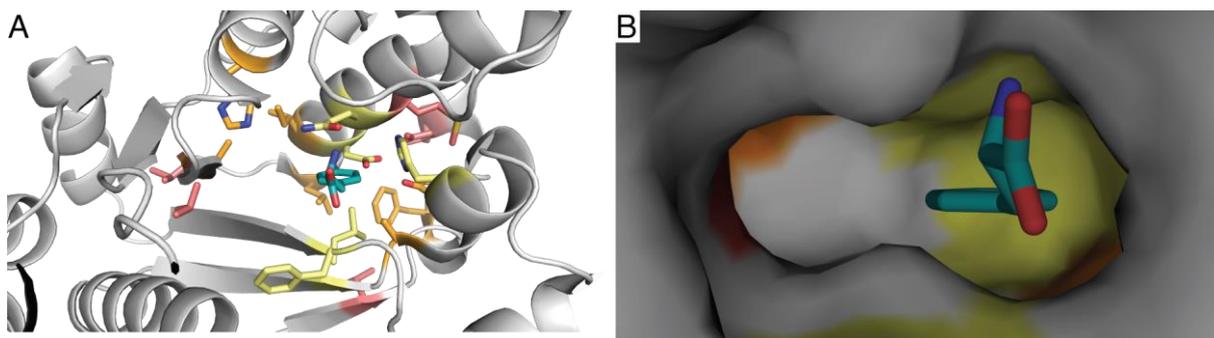


Abbildung 22: Kristallstruktur der *Mj*TyrRS (PDB ID: 1J1U), die als Grundlage für die Planung der SlonoMax Genbibliothek diente. (A) Cartoon-Darstellung des aktiven Zentrums der *Mj*TyrRS im Komplex mit Tyr (türkis). Die in der Zufallsbibliothek randomisierten Positionen sind in Stabform dargestellt. Die Seitenketten sind in Bezug auf den Anteil an mutierten Aminosäuren in der gesamten Bibliothek gefärbt: gleichmäßige Verteilung mit 5 % je Aminosäure (gelb); 50 % der Aminosäuren war die ursprüngliche Seitenkette und die verbliebenen 50 % randomisiert das eine Verteilung von 2,5 % je Aminosäure ergab (orange); der ursprüngliche Aminosäurerest war zu 90 % konserviert und die verbliebenen 10 % randomisiert das eine Verteilung von 0,5 % je Aminosäure ergab (rosa). (B) Darstellung der Proteinoberfläche um das aktive Zentrum. Farblich markiert ist der Grad der Randomisierung.

3.2.1.1 Selektion aus der SlonoMax Genbibliothek

Aus der erhaltenen SlonoMax Genbibliothek wurde 180 fmol, das entsprach $1 \cdot 10^{11}$ Genvarianten, die mittels PCR unter Verwendung der *Primer* ep1 und ep2 reamplifiziert (Abschnitt 2.3.3.5). Das PCR-Produkt wurde mit *Eam*I verdaut und anstelle der BpaRS auf den Selektionsvektor pSB4.6c2 kloniert. Nach der Transformation von elektrokompenten *E. coli* BL21 Zellen mit 20 μ g des Ligierungsproduktes wurde eine Transformanzahl von $1 \cdot 10^8$ Kolonien/ μ g eingesetzter Plasmid DNA ermittelt. Obwohl diese Transformanzahl für die Art der kompetenten Zellen verhältnismäßig niedrig ist, reichte die Gesamtzahl der erhaltenen

Transformanden angesichts der limitierten Kapazität des FACS-Verfahrens von ca. $1 \cdot 10^8$ sortierbaren Zellen für das geplante Experiment aus.

Für die Selektion der Varianten mit dem FACS-Verfahren wurden zuvor die auf Agarplatten kultivierten Transformanden in eine Flüssigkultur überführt, gemischt, in zwei Kulturen aufgeteilt und bei 30 °C und 180 rpm inkubiert. Dabei dient eine Kultur als Kontrolle, die während der exponentiellen Wachstumsphase ($OD_{550} \sim 0,3$), anstelle von 1 mM der nicht-natürlichen Aminosäure Hca, mit 1 mM der proteinogenen Aminosäure Tyr versetzt wurde. Eine Stunde nach der Supplementierung mit der (nicht-)natürlichen Aminosäure wurde die Expression der Aminoacyl-tRNA Synthetase-Varianten durch Zugabe von aTc (50 ng/ml) und nach einer weiteren Stunde, die Expression eGFP(39a) durch Zugabe von IPTG (500 mM) induziert. In Anschluss an Expression des Reporterproteins über Nacht (~ 16 h) wurde 1 ml Kultur ($\sim 1,5 \cdot 10^9$ *E. coli*-Zellen) geerntet und mit PBS gewaschen. Die Fluoreszenz des Reporterproteins wurde mit dem FACS-Gerät detektiert und die zu sortierenden Zellen, hinsichtlich gesteigerter (oder fehlender) Fluoreszenz, mittels Anpassung des Sortier-Fensters ausgewählt. Die sortierten Zellen wurden auf Agarplatten ausplattiert, über Nacht bei 37 °C inkubiert, bevor ein neuer Selektionszyklus begann (Abschnitt 2.4.1.2 und 2.6.3).

Ein Problem bei der Selektion einer mutierten Aminoacyl-tRNA Synthetase für den Einbau einer nicht-natürlichen Aminosäure besteht in der ursprünglichen Funktion der verwendeten *Mj*TyrRS. Bei fortwährender Selektion an *E. coli* Zellen, die eine eGFP(39a)-Fluoreszenz aufweisen (Positiv-Selektion), werden auch Varianten der Aminoacyl-tRNA Synthetase angereichert, die Tyr als Substrat erkennen. Zu erkennen ist dies, wenn die Fluoreszenzintensität der in Gegenwart von Tyr kultivierten Aminoacyl-tRNA Synthetase-Varianten im Verlauf der Selektionskampagne gleichermaßen oder sogar stärker steigt als die Fluoreszenzintensität der in Gegenwart von Hca kultivierten Aminoacyl-tRNA Synthetase-Varianten.

Um eine Anreicherung von solchen Aminoacyl-tRNA Synthetasen, die Tyr als Substrat erkennen, zu verhindern, wurde nach zwei Positiv-Selektionen eine Negativ-Selektion durchgeführt (Abbildung 23). Hierbei wurden aus den in Gegenwart von Tyr kultivierten Aminoacyl-tRNA Synthetase-Varianten jene Zellen mittels des Sortier-Fensters ausgewählt, bei denen kein Fluoreszenzsignal des Reporterproteins detektiert werden konnte. Im Anschluss an diese Negativ-Selektion wurden zur Steigerung der genetischen Diversität die zuvor sortierten Aminoacyl-tRNA Synthetase-Varianten einer *Error-Prone*-PCR unterworfen. Die darauffolgende Sequenzierung von 4 Klonen ergab 1 – 3 Mutationen, die eine Änderung einer Aminosäure zur Folge hatten. Auf diese Diversifizierung der Genbibliothek folgte eine erneute Positiv-Selektion. Nach insgesamt 5 Zyklen konnte jedoch kein Unterschied zwischen den in Gegenwart von Hca zu den in Gegenwart von Tyr kultivierten Aminoacyl-tRNA Synthetase-Varianten, bezogen auf die Fluoreszenz des Reporterproteins, detektiert werden.

Dies ließ vermuten, dass Aminoacyl-tRNA Synthetase-Varianten angereichert wurden, die die natürliche Aminosäure Tyr als Substrat erkennen (Abbildung 23, E und F).

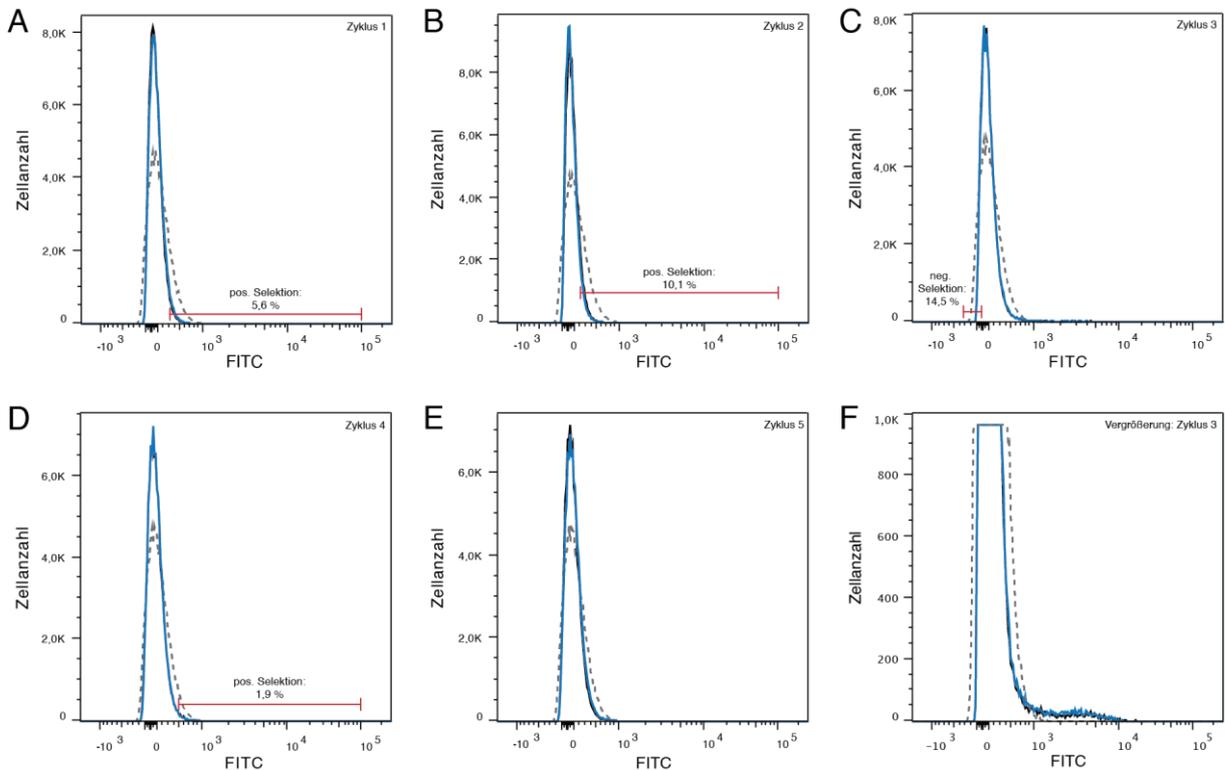


Abbildung 23: Selektionskampagne der SlonoMax Bibliothek. Eine auf der *MjTyrRS* basierenden Genbibliothek wurde bezüglich der Fähigkeit, Hca zu erkennen, selektiert. Dargestellt ist das Fluoreszenzsignal von eGFP(39a), erzeugt durch die Amber-Stop Codon Suppression in Gegenwart von 1 mM Hca (blau) oder 1 mM Tyr (schwarz) nach 16-stündiger Expression des Reporterproteins. In Grau ist das Fluoreszenzsignal der nicht transfizierten, nativen, *E. coli* BL21 Zellen dargestellt. Plots A – E zeigt die Selektionszyklen 1 – 5, in Plot F ist eine vergrößerte Ansicht des dritten Zyklus (C) dargestellt. Die sortierte Zellpopulation ist mit einem roten Balken gekennzeichnet und darüber ist der prozentuelle Anteil dieser Zellpopulation vermerkt. Für die Messung wurden die Spannung der Laser am FACS-Gerät wie folgt eingestellt: FITC = 700 V; FSC = 400 V; SSC = 550 V.

Die molekulare Erkennung von Tyr durch die *MjTyrRS* erfolgt vor allem mit den die Seitenketten Tyr³² und Asp¹⁵⁸ im aktiven Zentrum, wo Wasserstoffbrückenbindungen von den Aminosäureseitenketten zu der Hydroxylgruppe der Tyr-Seitenkette ausgebildet werden. Die Orientierung der Seitenkette Asp¹⁵⁸ wird durch eine weitere Wasserstoffbrückenbindung mit Gln¹⁰⁹ stabilisiert (Kobayashi *et al.*, 2003).

Für die Bildung einer Carbonsäure-Phosphorsäure-Anhydrid-Bindung zwischen der Substrat-Aminosäure und ATP wird Tyr unter anderem durch Wasserstoffbrückenbindungen zwischen der α -Aminogruppe als auch der Carbonsäuregruppe und der Seitenkette von Gln¹⁷³ in Position gehalten. Ausrichtung der Seitenkette von Gln¹⁷³ wird wiederum von Gln¹⁵⁵ durch eine Wasserstoffbrückenbindung zwischen den Seitenketten stabilisiert (Kobayashi *et al.*, 2003). Das Aminosäuresequenz-Alignment von 4 Klonen und der publizierten *MjTyrRS* bestärkte die Vermutung, dass vorallem Aminoacyl-tRNA Synthetase-Varianten angereichert wurden, die

Tyr als Substrat erkennen, denn bei allen 4 selektierten Aminoacyl-tRNA Synthetasen wurden die Aminosäuren Gln¹⁵⁵ und Asp¹⁵⁸ detektiert. Die Aminosäure Gln¹⁷³ wurde in dieser Bibliothek nicht randomisiert und blieb daher unverändert. Zudem wurde bei 2 der selektierten Aminoacyl-tRNA Synthetasen auch Gln¹⁰⁹ nachgewiesen (Abbildung 24). An der Aminosäure-Position 32 hingegen wurde keine Homologie festgestellt.

Abbildung 24: Aminosäuresequenz-Alignment von 4 nach der Selektionskampagne gepickten Einzelklonen und MjTyrRS (UniProt ID: Q57834). Dargestellt ist die Aminosäuresequenz des gesamten aktiven Zentrums (210 von insgesamt 306 Aminosäuren), die durch eine Sequenzierung abgedeckt wurden. Randomisierte Aminosäurepositionen der SlonoMax-

	1	10	20	30	40	50	60	70
MjTyrRS	MDEFEmIKRNTSEIISEEELREVLKKDEKSA	YIGFEP	SGKIHLGHYLQIKKMIDLQ	NAGFDII	ILLADLH			
Bibliothek1.....X.....X.X.....
Klon11.....T.....C.....P
Klon21.....C.....I.....T
Klon31.....M.....V.....
Klon41.....V.....D.....T
		80	90	100	110	120	130	140
MjTyrRS	AYLNQK	GELDEIRKIGDYNK	KVFEAMGLKAKYVYGSEF	QLDKDYTLNVYRLAL	KTTLKRARRSMELIARE			
BibliothekX.....XX.....X.....
Klon1L.....
Klon2HG.....
Klon3V.....
Klon4
		150	160	170	180	190	200	210
MjTyrRS	DENPKVAEVIYPIMQVNDIHYLGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGLDGEGKMSSSKG							
BibliothekXX.....XX.....X.....XXX.....X.....
Klon1M.....
Klon2M.....V.....Q.....
Klon3L.....G.....
Klon4Q.....

Bibliothek sind mit dem Buchstaben x markiert, dabei kennzeichnen die Farben Gelb (5 %), Orange (2,6%) und Rot (0,5 %) den Randomisierungsgrad. Kleinbuchstaben kennzeichnen einen gezielten Aminosäureaustausch, um ein verkürztes Protein durch ein zweites Start-Methionin zu verhindern (Schiefner & Skerra, 2015).

3.2.1.2 Selektion aus der SlonoMax Genbibliothek nach einer Vorselektion

Eine Genbibliothek mit einer errechneten Komplexität von $1,3 \cdot 10^{22}$ Genvarianten enthält erwartungsgemäß viele Varianten, die aufgrund von Mutationen funktionsunfähig sind. Dadurch wird die Selektion einer geeigneten Aminoacyl-tRNA Synthetase mit dem FACS-Verfahren erschwert. Daher erschien es von Vorteil, funktionsfähige Aminoacyl-tRNA Synthetasen, die vorzugsweise eine nicht-natürliche Aminosäure erkennen, in einer Vorselektion auf Agarplatten anzureichern.

Diese Vorselektion basiert auf der Suppression eines Amber-Stop Codons, welches in die Sequenz eines Antibiotikaresistenzgens eingeführt wurde. Das dafür verwendete Plasmid trägt zusätzlich zum Ampicillin-Resistenz vermittelnden Gen der β -Lactamase (ampR) das Chloramphenicol-Resistenz vermittelnde Gen der Chloramphenicol-Acetyltransferase (camR) mit dem Amber-Stopp-Codon an der Aminosäureposition 112 (camR(a112)) (Pastrnak *et al.*,

2000) als Selektionsmarker. Das Chloramphenicol-Acetyltransferasegen wurde gewählt, da das daraus resultierende Enzym im Vergleich zu β -Lactamasen nicht in das Medium sezerniert wird. Das stellt sicher, dass der Selektionsdruck auch bei langer Inkubationszeit aufrechterhalten bleibt und die Bakterienzelle nur bei einer erfolgreichen Suppression des Amber-Stop Codons in der Chloramphenicol-Acetyltransferase-Sequenz überlebt.

Ausgehend von den Plasmiden pSB15.CafRS#30d53 (erhalten von Andreas Reichert, Lehrstuhl für Biologische Chemie) und pSB4.6c2 (Abbildung 25 A) wurde das Plasmid pSB16.6c2 (Abbildung 25 B) konstruiert, das für die Vorselektion auf Nährmediumsplatten und anschließend auch für die Selektion mit dem FACS-Verfahren verwendet werden konnte. Das Plasmid kodiert demnach die für die Vorselektion benötigte Gene *camR*(112a) als transkriptionelle Fusion mit *ampR* und dem *tet*-Repressor. Des Weiteren befindet sich die Aminoacyl-tRNA Synthetase-Sequenz (hier *BpaRS*) unter der Kontrolle des konstitutiven *proS*-Promotors, die Suppressor-tRNA unter Kontrolle des konstitutiven *lpp*-Promotors und das für die Selektion mittels FACS benötigte Reporterprotein *eGFP*(39a) unter der Kontrolle des mit *aTc* induzierbaren *tetA* Promotor/Operator-Systems (Skerra, 1994).

Bevor das klonierte Plasmid für die Vorselektion der SlonoMax-Genbibliothek verwendet werden konnte, wurde der optimale Selektionsdruck, vermittelt durch die Chloramphenicol-Konzentration im Agarosemedium, evaluiert. Zu diesem Zweck wurde eine bereits charakterisierte Aminoacyl-tRNA Synthetase, die *BpaRS*, ausgewählt, welche nachweislich *Bpa* erkennt und diese nicht-natürliche Aminosäure in die Polypeptidkette einbaut. Aus einer Flüssigkultur mit *E. coli* BL21 Zellen, die das Plasmid pSB16.6c2 trugen, wurden Bakterienzellen entnommen und auf Selektionsplatten, bestehend aus LB-Agar mit 1 mM *Bpa*, 100 μ g/ml Ampicillin und 0, 10, 15, 20 oder 25 μ g/ml Chloramphenicol, ausplattiert. Nach 22-stündiger Inkubation bei 37 °C waren Kolonien sichtbar, die ausgezählt werden konnten.

Tabelle 1: Evaluierung der optimalen Chloramphenicol-Konzentration zur Selektion auf die Suppression des Amber-Stop Codons. Angeführt ist die Zahl der *E. coli* BL21 Kolonien mit pSB16.6c2, die auf Agarplatten mit 1 mM *Bpa*, 100 μ g/ml Ampicillin und der genannten Chloramphenicol-Konzentration nach 12 h oder 22 h Inkubation gezählt werden konnten.

Chloramphenicol [μ g/ml]	Kolonien	Inkubationsbedingungen
0	$3,5 \cdot 10^8$	12 h, 37 °C
10	$3,2 \cdot 10^8$	22 h, 37 °C
15	$2,1 \cdot 10^8$	22 h, 37 °C
20	$1,1 \cdot 10^8$	22 h, 37 °C
25	$1,6 \cdot 10^7$	22 h, 37 °C

Dabei wurde eine Abnahme der Kolonienanzahl bei steigender Chloramphenicol-Konzentration im Bereich von 15 – 20 μ g/ml festgestellt. Da von einer Aminoacyl-tRNA Synthetase-Variante aus einer nativen Bibliothek nicht zu erwarten ist, dass diese eine

nicht-natürliche Aminosäure ähnlich gut erkennt wie eine bereits optimierte Aminoacyl-tRNA Synthetase, erschien ein moderaterer Selektionsdruck von 15 $\mu\text{g/ml}$ im Vergleich zu 20 $\mu\text{g/ml}$ Chloramphenicol sinnvoll (Tabelle 1).

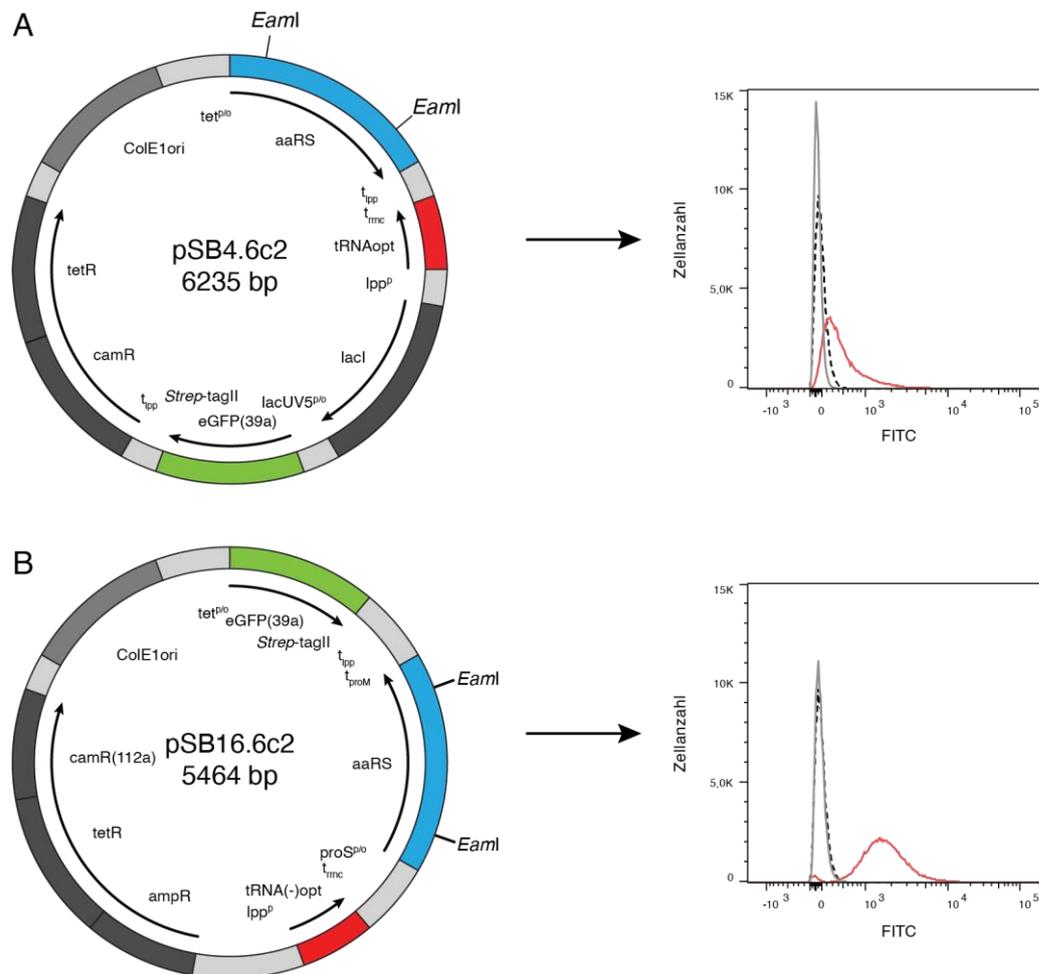


Abbildung 25: Vergleich der Fluoreszenzintensität durch die Suppression von eGFP(39a) mit den Plasmiden pSB4.6c2 oder pSB16.6c2. (A) Schematische Darstellung des Expressionsplasmids pSB4.6c2 und des aus der Expression von eGFP(39a) resultierenden Fluoreszenzsignals bei der FACS-Analyse in Gegenwart von 1 mM Bpa (rot) oder 1 mM Tyr (grau). Auf dem Plasmid wird die Sequenz des Reporterproteins eGFP(39a) inklusive des *Strep*-tag II (Skerra & Schmidt, 2000) am 3'-Ende unter der Kontrolle des mit IPTG induzierbaren lacUV5 Promotor/lac-Operators (lacUV5^{pl}) (Silverstone *et al.*, 1970) kodiert, gefolgt von dem lpp-Terminator (t_{pp}). Zusätzlich trägt das Plasmid die Gensequenz für die BpaRS (Brustad *et al.*, 2008) unter der Kontrolle des mit aTc regulierbaren tetA Promotor/Operators (tet^{pl}) (Skerra, 1994), gefolgt vom lpp-Terminator (t_{pp}) und der dazugehörigen tRNA (tRNA(-)opt) (Wang *et al.*, 2001) unter der Kontrolle des lpp-Promotors (lpp^P), welche vom rnc-Terminator (t_{rnc}) terminiert wird. Weitere Elemente auf dem Plasmid sind der ColE1 Replikationsursprung (ColE1 ori) und der LacI-Repressor (lacI). Die konstitutive Expression der Transkriptionsfusion dem Gen der Chloramphenicol-Acetyltransferase (camR) und dem tet-Repressor (tetR) ermöglichte die effektive Repression des Promotors für die Expression der BpaRS wie auch die antibiotische Selektion. (B) Schematische Darstellung des Expressionsplasmids pSB16.6c2 und des aus der Expression von eGFP(39a) resultierenden Fluoreszenzsignals. Anders als bei pSB4.6c2 ist eGFP(39a), inklusive des *Strep*-tag II (Skerra & Schmidt, 2000) am 3'-Ende, unter der Kontrolle des mit aTc regulierbaren tetA Promotor/Operators (tet^{pl}) (Skerra, 1994), ausgetauscht. Des weiteren ist auf diesem Plasmid die BpaRS (Brustad *et al.*, 2008) unter der Kontrolle des konstitutiven proS-Promotors kodiert, gefolgt vom proM-Terminator. Die konstitutive Expression der Transkriptionsfusion aus dem Ampicillin-Resistenzgen (AmpR), tet-Repressor (tetR) und dem Chloramphenicol-Resistenzgen mit einem Amber-Stop Codon an Position 112 ermöglichte sowohl die effektive Repression des Promotors für die Expression von eGFP(39a) wie auch eine generelle (Amp) und eine suppressionsabhängige (Cam) antibiotische Selektion.

Bei dem Plasmid pSB16.6c2 wurden im Vergleich zum ursprünglichen Plasmid, pSB4.6c2, regulatorische Elemente verändert, die sich auch auf die Expression beziehungsweise die Suppression des Amber-Stop Codons des Reporterproteins auswirken könnten. In einem FACS-Parallelversuch ergab die Expression von eGFP(39a) in den *E. coli* BL21 Zellen mit pSB16.6c2 eine um den Faktor 3 höhere Fluoreszenzintensität als bei Verwendung des Plasmids pSB4.6c2 (Abbildung 25). Das weist darauf hin, dass eine konstitutive Expression der BpaRS im Gegensatz zu einer induzierten Expression in der exponentiellen Wachstumsphase von Vorteil ist. Da bei der konstitutiven Expression der Aminoacyl-tRNA Synthetase auch während der Wachstumsphase die Termination an endogenen Amber-Stop Codons durch den Einbau der nicht-natürlichen Aminosäure unterdrückt wird, kann dieser Promotor jedoch nur in Kombination mit einer nicht-natürlichen Aminosäure verwendet werden, deren Einbau für die Bakterienzelle nicht toxisch ist.

Für die Vorselektion auf der Agarplatte und die anschließende Selektionskampagne mit dem FACS-Verfahren wurden die reamplifizierten und mit *EamI* verdauten Gensequenzen der SlonoMax-Genbibliothek anstelle der BpaRS auf den Vektor pSB16.6c2 kloniert. Mit dem Ligierungsprodukt wurden elektrokompente *E. coli* BL21 Zellen transformiert und auf Selektionsplatten (100 µg/ml Ampicillin und 15 µg/ml Chloramphenicol) mit 1 mM Hca oder 1 mM Tyr ausplattiert (Abschnitt 2.4.1.3). Nach 22 h Inkubation bei 37 °C wurden auf den LB-Selektionsplatten mit 1 mM Hca $4,1 \cdot 10^4$ Kolonien und auf den LB-Selektionsplatten mit 1 mM Tyr $6,8 \cdot 10^4$ Kolonien/µg eingesetzter Plasmid DNA gezählt. Mit den elektrokompenten *E. coli* BL21 Zellen wurde auch ein Transformationsexperiment durchgeführt, um die Transformationseffizienz der Zellen ohne Selektionsdruck durch die Suppression des Amber-Stop Codons feststellen zu können. Dafür wurden die Zellen mit pSB16.6c2 transformiert und auf LB-Agarplatten mit 100 µg/ml Ampicillin ausplattiert. Nach 12-stündiger Inkubation bei 37 °C wurden $9,5 \cdot 10^7$ Transformanden/µg eingesetzter Plasmid DNA gezählt. Dieses Verhältnis der Transformanden auf den Selektionsplatten zu den Transformanden auf den normalen Agarplatten wies darauf hin, dass nur ein geringer Anteil der Varianten in der Lage ist, Hca oder Tyr als Substrat zu erkennen und bestätigt damit den Nutzen einer Vorselektion bei hochkomplexen Bibliotheken.

Für die weitere Selektion der erhaltenen $4,1 \cdot 10^4$ Aminoacyl-tRNA Synthetase-Varianten mit dem FACS-Verfahren wurden die auf Selektionsplatten mit Hca kultivierten Transformanden in eine Flüssigkulturen überführt und bei 30 °C und 180 rpm inkubiert. Während der exponentiellen Wachstumsphase ($OD_{550} \sim 0,4$) wurde die Flüssigkultur in zwei Kulturen geteilt und entweder mit 1 mM Hca oder mit 1 mM Tyr versetzt. Nach einer Stunde wurde die Expression des Reporterproteins eGFP(39a) durch Zugabe von aTc (200 ng/ml) induziert und nach weiteren 16 h wurde je 1 ml Kultur für die Selektion geerntet. Die eGFP(39a)-Fluoreszenz der mit Hca wie auch der mit Tyr versetzten Kultur wurde mit dem FACS-Gerät detektiert und

verglichen. Basierend auf den daraus resultierenden Unterschieden wurden die zu sortierenden Zellen mittels Anpassung des Sortier-Fensters ausgewählt.

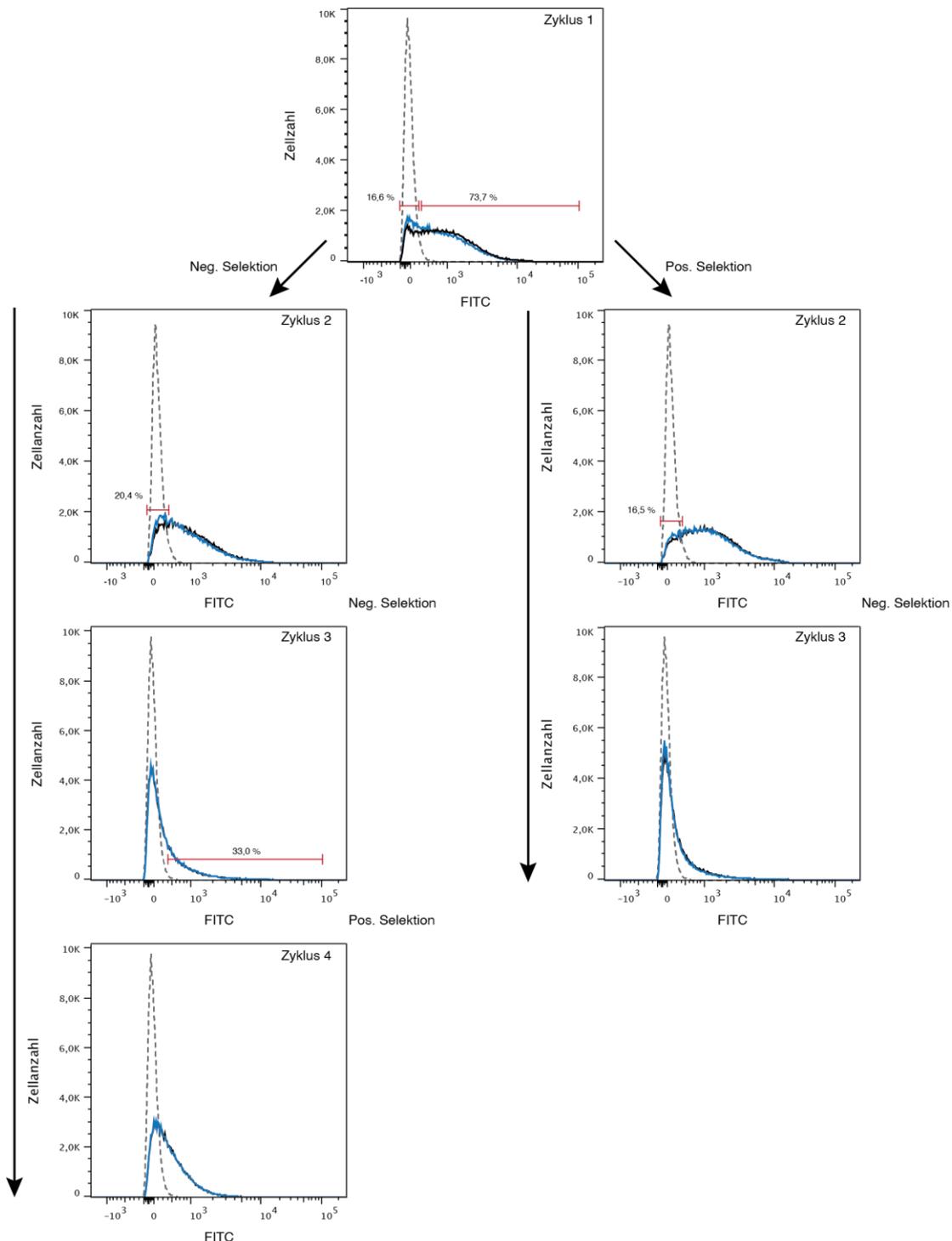


Abbildung 26: Selektionskampagnen mit der SlonoMax-Bibliothek nach Vorselektion auf Agarplatten. Eine auf der *MjTyrRS* basierende Genbibliothek wurde bezüglich der Fähigkeit, Hca zu erkennen, selektiert. Dargestellt ist das Fluoreszenzsignal von eGFP(39a), erzeugt durch die Amber-Stop Codon Suppression in Gegenwart von 1 mM Hca (blau) und 1 mM Tyr (grau) nach 16-stündiger Expression des Reporterproteins. In Schwarz ist das Fluoreszenzsignal der nicht transfizierten, nativen, *E. coli* BL21 Zellen dargestellt. Ausgehend von der Vorselektion wurden zwei verschiedene Strategien angewandt (linke und rechte Spalte). Die sortierten Zellen sind mit einem roten Balken gekennzeichnet und der prozentuale Anteil dieser bezieht sich auf alle im aktuellen Zyklus analysierten Zellen. Für die Messung wurden die Spannung der Laser am FACS-Gerät wie folgt eingestellt: FITC = 650 V; FSC = 300 V; SSC = 500 V.

Die sortierten Zellen wurden auf Agarplatten ausplattiert und über Nacht bei 37 °C inkubiert, bevor sie für einen neuen Zyklus erneut in eine Flüssigkultur eingebracht wurden (Abschnitt 2.4.1.2 und 2.6.3). Im Vergleich zu der Selektion ohne Vorselektion auf Agarplatten (Abschnitt 3.2.1.1, Abbildung 23) war bei der Detektion des Reporterproteins eine Verschiebung des Fluoreszenzsignals nach rechts zu beobachten, was auf die Anreicherung an Aminoacyl-tRNA Synthetasen, die eine (nicht-)natürliche Aminosäure als Substrat erkennen können, hindeutete. Ein Unterschied zwischen der *E. coli* BL21-Kultur der Genvarianten kultiviert mit Hca und der genetisch identischen Kultur mit Tyr als Kontrolle konnte nicht festgestellt werden (Abbildung 26). Daher wurden ausgehend davon zwei unterschiedliche Sortier-Strategien parallel verfolgt. In einer der beiden Strategien wurde nach der positiven Vorselektion eine erneute Positiv-Selektion im zweiten Zyklus, gefolgt von einer Negativ-Selektion im dritten Zyklus, durchgeführt. Weder im zweiten noch im dritten Zyklus konnte ein Unterschied zwischen der mit Hca oder mit Tyr versetzten *E. coli* BL21-Kultur festgestellt werden (Abbildung 26, rechte Spalte). Bei der anderen Strategie wurden ausgehend von der positiven Vorselektion zwei aufeinanderfolgende Negativ-Selektionen durchgeführt, um Aminoacyl-tRNA Synthetasen, die Tyr als Substrat erkennen, abzureichern. Danach zeigte die *E. coli* BL21-Kultur der Genvarianten in Gegenwart von Hca ein minimal besseres Fluoreszenzsignal (1 %, ausschließlich erkennbar mit der FACS-Software) als die Kontrollkultur in Gegenwart von Tyr bezogen auf das gleiche Sortier-Fenster. Jedoch ließ sich diese Tendenz nach einer darauffolgenden Positiv-Selektion nicht bestätigen (Abbildung 26, linke Spalte). Das bedeutete, dass die Selektion einer Aminoacyl-tRNA Synthetase, die Hca als Substrat erkennt, mit der SlonoMax-Genbibliothek nicht erfolgreich war.

Eine Sequenzierung (Abbildung 27) von je 4 gepickten Klonen nach den beiden Selektionskampagnen gab einen Einblick in das Mutationsmuster der selektierten Bibliothek, aber es konnten im Vergleich zu der Sequenzanalyse in Abschnitt 3.2.1.1 keine klaren Schlüsse gezogen werden.

A

	1	10	20	30	40	50	60	70		
<i>MjTyrRS</i>	MDEFEM	IKRNTSEI	I	SEEELREVL	KKDEKSA	YIGFEP	SGKIHLGHYL	QIKKMIDLQ	NAGFDII	ILLADLH
Bibliothek	l.....	x.....	x.x.x	x
Klon1	l.....	T.....	R.....P
Klon2	l.....	T.....	R.....L
Klon3	l.....	H.....	A.....M
Klon4	l.....	H.....	T.....Q

	80	90	100	110	120	130	140
<i>MjTyrRS</i>	AYLNQKGELDE	IRKIGDYNK	KVFEAMGLK	KAKYVYGSEF	QLDKDYTLN	VYRLALKTT	LKRARRSMELIARE
Bibliothek	x.....	xx.....	x.....
Klon1	H.....
Klon2	M.....
Klon3	A.....
Klon4	A.....

	150	160	170	180	190	200	210
<i>MjTyrRS</i>	DENPKVAEVI	YPIMQVND	IHYLGVDVAV	GGMEQRKIH	MLARELLPK	KVVCIHNPV	LTLGDLGEGKMSSSKG
Bibliothek	xx.....	xx.....	x.....	xxx.....	x.....
Klon1
Klon2	A.....	SF.....	G.....	W.....
Klon3	E.....	W.....	C.....
Klon4	E.....	W.....	C.....

B

	1	10	20	30	40	50	60	70		
<i>MjTyrRS</i>	MDEFEM	IKRNTSEI	I	SEEELREVL	KKDEKSA	YIGFEP	SGKIHLGHYL	QIKKMIDLQ	NAGFDII	ILLADLH
Bibliothek	l.....	x.....	x.x.x	x
Klon1	l.....	T.....	R.....P
Klon2	l.....	M.....	P.....L
Klon3	l.....	S.....	A.....M
Klon4	l.....	L.....	C.....Q

	80	90	100	110	120	130	140
<i>MjTyrRS</i>	AYLNQKGELDE	IRKIGDYNK	KVFEAMGLK	KAKYVYGSEF	QLDKDYTLN	VYRLALKTT	LKRARRSMELIARE
Bibliothek	x.....	xx.....	x.....
Klon1	N.....	A.....
Klon2	G.....
Klon3
Klon4	L.....

	150	160	170	180	190	200	210
<i>MjTyrRS</i>	DENPKVAEVI	YPIMQVND	IHYLGVDVAV	GGMEQRKIH	MLARELLPK	KVVCIHNPV	LTLGDLGEGKMSSSKG
Bibliothek	xx.....	xx.....	x.....	xxx.....	x.....
Klon1	R.....	FM.....	T.....
Klon2	L.....	V.....	P.....	Y.....	C.....
Klon3	S.....	E.....
Klon4	L.....	LP.....	P.....

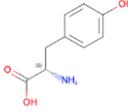
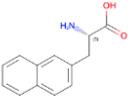
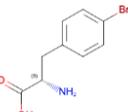
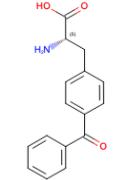
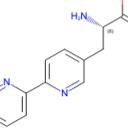
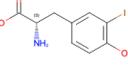
Abbildung 27: Aminosäuresequenz-Alignment von je 4 nach zwei Selektionskampagnen gepickten Einzelkolonien im Vergleich zu *MjTyrRS*. Dargestellt ist die Aminosäuresequenz der zentralen Enzymregion (210 von insgesamt 306 Aminosäuren), die durch eine Sequenzierung abgedeckt werden konnte. Randomisierte Aminosäurepositionen der SlonoMax-Bibliothek sind mit dem Buchstaben x markiert, dabei kennzeichnen die Farben Gelb (5 %), Orange (2,6 %) und Rot (0,5 %) den Randomisierungsgrad. Kleinbuchstaben kennzeichnen den gezielten Aminosäureaustausch um ein verkürztes Protein durch ein zweites Start-Methionin zu verhindern. (A) Sequenzen aus der Selektionskampagne, die mit einer Negativ-Selektion in Runde 1 startete, siehe Abbildung 26, linke Spalte und (B) die mit einer Positiv-Selektion in Runde 1 startete, siehe Abbildung 26, rechte Spalte.

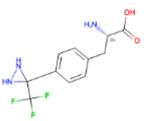
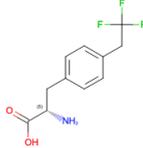
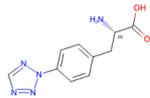
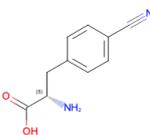
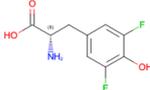
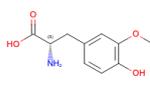
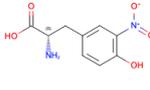
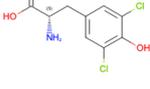
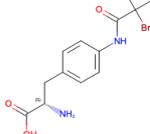
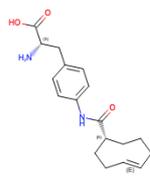
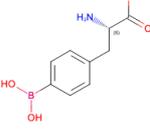
3.2.2 Selektion einer Aminoacyl-tRNA Synthetase einer auf Substratgruppen fokussierten Zufalls-Bibliothek der *Mj*TyrRS

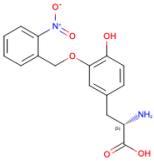
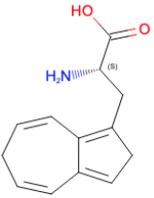
3.2.2.1 Strukturelle Analyse evolvierter Aminoacyl-tRNA Synthetasen

Die strukturelle Überlagerung (Tabelle 2) von den 18 Aminoacyl-tRNA Synthetasen, die ausgehend von der *Mj*TyrRS für die Erkennung einer nicht-natürlichen Aminosäure evolviert wurden, gab Einblicke in die Orientierung der nicht-natürlichen Aminosäuren im aktiven Zentrum bei der Substraterkennung.

Tabelle 2: Für die strukturelle Überlagerung verwendeten Aminoacyl-tRNA Synthetasen. Aufgelistet sind die PDB IDs, die für die Überlagerung verwendete Protein-Kette, das Aminosäure-Substrat der jeweiligen Aminoacyl-tRNA Synthetase und der in der Kristallstruktur erkennbare Ligand. In Rot sind Aminoacyl-tRNA Synthetasen mit sterisch anspruchsvollen Substraten markiert, welche für Abbildung 28 verwendet wurden.

PDB	Titel in der Protein Date Bank	Kette	Aminosäure	Ligand	Struktur
1J1U	Crystal structure of archaeal tyrosyl-tRNA synthetase complexed with tRNA(Tyr) and L-tyrosine	A	Tyrosin	Tyr + tRNA	
1ZH6	Crystal Structure of L-3-(2-naphthyl)alanine-tRNA synthetase in complex with L-3-(2-naphthyl)alanine	A	Beta-(2-naphthyl)-Alanin	NAL	
2AG6	Crystal structure of p-bromo-l-phenylalanine-tRNA synthetase in complex with p-bromo-l-phenylalanine	A	4-Acetyl-L-Phenylalanin	4AF	
2HGZ	Crystal structure of a p-benzoyl-L-phenylalanyl-tRNA synthetase	A	p-Banzoyl-Phenylalanin	PBF	
2PXH	Crystal structure of a bipyridylalanyl-tRNA synthetase	A	3-(2,2'-Bipyridin-5-yl)-L-Alanin	BP5	
2ZP1	Structural basis of iodo-tyrosine recognition by engineered archeal tyrosyl-tRNA synthetase	A	3-Iodo-Tyrosin	IYR	

3D6U	Crystal structure of 4-(trifluoromethyldiazirinyl)phenylalanyl-tRNA synthetase	A	4-[3-(Trifluoromethyl)diaziridin-3-yl]-L-Phenylalanin	T11	
3D6V	Crystal structure of 4-(trifluoromethyldiazirinyl)phenylalanyl-tRNA synthetase	A	4-(2,2,2-Trifluoroethyl)-L-Phenylalanin	TFQ	
3N2Y	Crystal structure of tyrosyl-tRNA synthetase complexed with p-(2-tetrazolyl)-phenylalanine	A	4(2H-Tetrazol-2-yl)-L-Phenylalanin	TEF	
3QE4	An evolved aminoacyl-tRNA Synthetase with atypical polysubstrate specificity	A	4-Cyano-L-Phenylalanin	4CF	
4HJX	Crystal structure of F2YRS complexed with F2Y	A	3,5-Difluoro-L-Tyrosin	F2Y	
4HPW	Crystal structure of Tyrosine-tRNA ligase mutant complexed with unnatural amino acid 3-o-methyl-Tyrosine	A	3-Metgoxy-L-Tyrosin	3YM	
4NDA	Crystal structure of 3-nitro-tyrosine tRNA synthetase (5B) bound to 3-nitro-tyrosine	A	3-Nitro-Tyrosin	NIY	
4NX2	Crystal structure of DCYRS complexed with DCY	A	3,5-Dichloro-L-Tyrosin	2LT	
4PBS	Crystal structure of the M. jannaschii F9 tRNA synthetase variant bound to 4-(2-bromoisobutyramido)-phenylalanine (BibaF)	A	4-[(2-Bromo-2-methylprpanoyl)amino]-L-Phenylalanin	2L7	
4PBT	Crystal structure of the M. jannaschii G2 tRNA synthetase variant bound to 4-trans-cyclooctene-amidopheylalanine (Tco-amF)	A	4{[81R,4E)-Cyclooct-4-en-1-ylcarbonyl]amino}-L-Phenylalanin	2LQ	
5N5U	Structure of p-boronophenylalanyl tRNA synthetase in complex with p-boronophenylalanine and adenosine monophosphate	A	4-Borono-L-Phenylalanin	7N8 + AMP	

5L7P	In silico-powered specific incorporation of photocaged Dopa at multiple protein sites	B	2~{S}-2-azanyl-3-[3-[(2-nitrophenyl)methoxy]-4-oxidanyl-phenyl] Propansäure	BUU	
5NSF	Structure of AzuAla	B	(2~{S})-2-Azanyl-3-(2,6-dihydroazulen-1-yl) Propansäure	96Z	

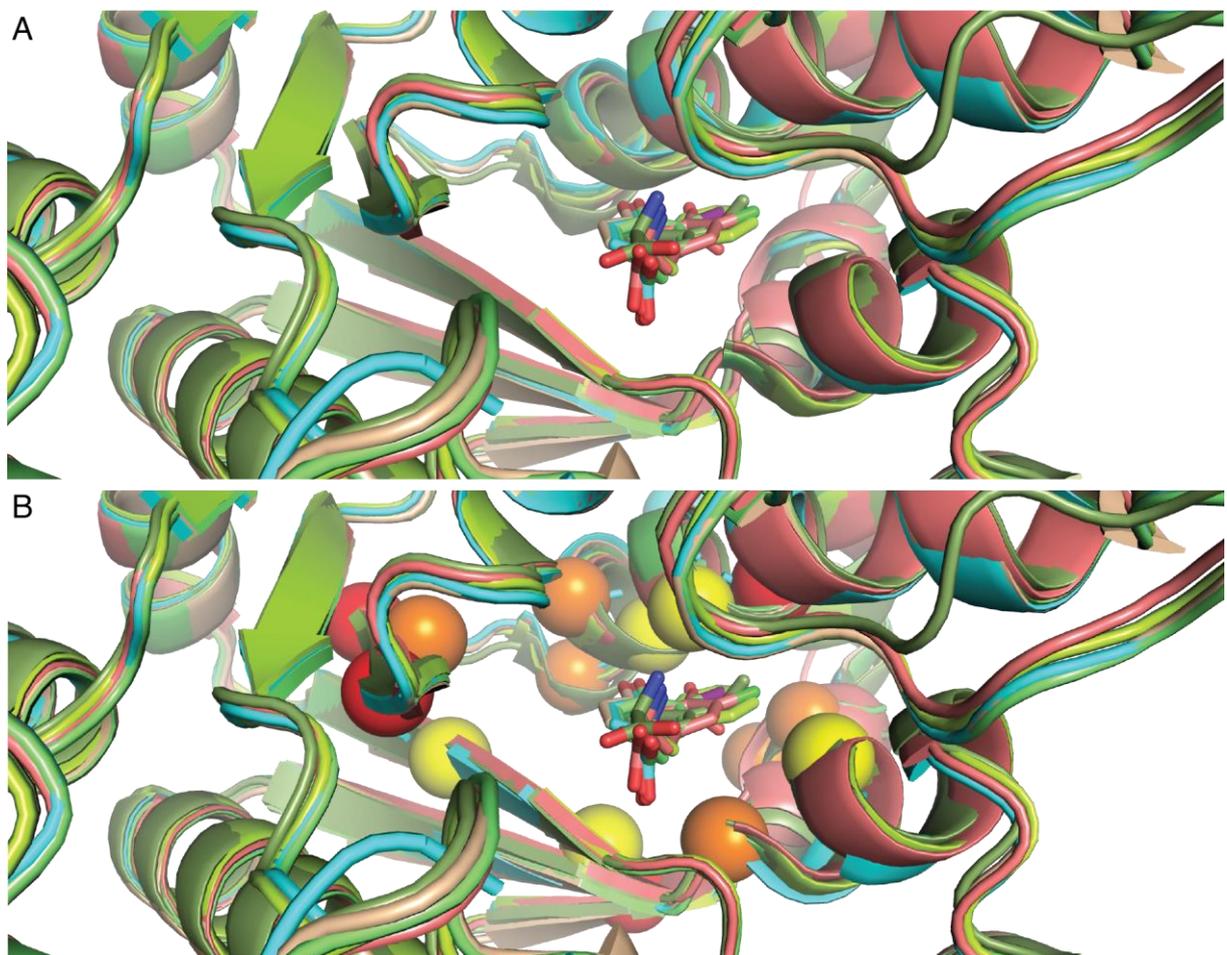


Abbildung 28: Überlagerung der Kristallstrukturen von Aminoacyl-tRNA Synthetasen, die ausgehend von der *MjTyrRS* für die Erkennung von β -verzweigten nicht-natürlichen Aminosäure engineered wurden. (A) Überlagert wurden die *MjTyrRS* und Aminoacyl-tRNA Synthetasen, die die nicht-natürlichen Aminosäuren 3-Iod-Tyr (PDB ID: 2ZP1), 3-Methyl-L-Tyr (PDB ID: 4HPW), 3-Nitro-Tyr (PDB ID: 4NDA), AzAla (PDB ID: 5NSF) und 3,5-Dichlor-L-Tyr (PDB ID: 4NX2) erkennen. (B) Durch Kugeln hervorgehoben sind die bei der SlonoMax Bibliothek randomisierten Aminosäurepositionen. Die Farbe beschreibt dabei die Verteilung der randomisierten Aminosäuren, Gelb (5 %), Orange (2,5 %) und Rot (0,5 %).

Dabei fiel auf, dass die Kavität in der Nähe des aktiven Zentrums nicht zur Bindung der nicht-natürlichen Aminosäure genutzt wurde. Unklar ist, ob dieser Raum im Verlauf der natürlichen Bindung beziehungsweise während der Adenylierung oder der Beladung der tRNA benötigt wird.

Durch gerichtete Evolution wurde Platz im aktiven Zentrum für voluminöse nicht-natürlichen Aminosäuren geschaffen (Abbildung 28). Diese nicht-natürlichen Aminosäuren weisen alle eine Ringstruktur am $\text{C}\beta$ auf, die in *para* und *meta* Position substituiert sind. Sie ähneln der nicht-natürlichen Aminosäure Hca, strukturell ebenfalls eine 6-Ringstruktur am $\text{C}\beta$, an der aber in *meta*- und *ortho*-Position ein weiterer Aromat anelliert ist (Abbildung 29). Damit wurde klar, dass die zur Randomisierung gewählten Positionen in der SlonoMax Bibliothek (Abbildung 28B) den zu mutierenden Bereich nicht in idealer Weise abdecken, um daraus eine Aminoacyl-tRNA Synthetase für die Erkennung von Hca zu selektieren (Abbildung 22).

Bei dem strukturellen Vergleich der Aminoacyl-tRNA Synthetasen fielen zwei Vertreter besonders auf, die AzAlaRS (PDB ID: 5NSF) und die PhotoDopaRS (PDB ID: 5L7P). Die AzAlaRS (Baumann *et al.*, 2019b) erkennt Azulenalalin (AzAla). Bei dieser nicht-natürlichen Aminosäure befindet sich an $\text{C}\beta$ ein 5-Ring, an den ein 7-Ring anelliert ist (Abbildung 29). Die PhotoDopaRS (Hauf *et al.*, 2017) erkennt *ortho*-Nitrobenzyl Dopamin (PhotoDopa), eine nicht-natürliche Aminosäure, die durch ihre photosensitive Schutzgruppe sterisch besonders anspruchsvoll ist (Abbildung 29). In der Literatur ist zudem eine Aminoacyl-tRNA Synthetase (HqAlaRS) beschrieben (Liu *et al.*, 2013), deren Struktur zwar noch nicht gelöst ist, die aber mit 8-Hydroxychinolin-Alanin (HqAla) eine nicht-natürliche Aminosäure erkennt, die Hca strukturell ähnelt (Abbildung 29).

Oftmals erkennen für nicht-natürliche Aminosäuren engierte Aminoacyl-tRNA Synthetasen auch andere, strukturell ähnliche nicht-natürliche Aminosäuren. Eine solche Promiskuität konnte jedoch bei den Aminoacyl-tRNA Synthetasen HqAlaRS, AzAlaRS und PhotoDopaRS mit der nicht-natürlichen Aminosäure Hca nicht beobachtet werden.

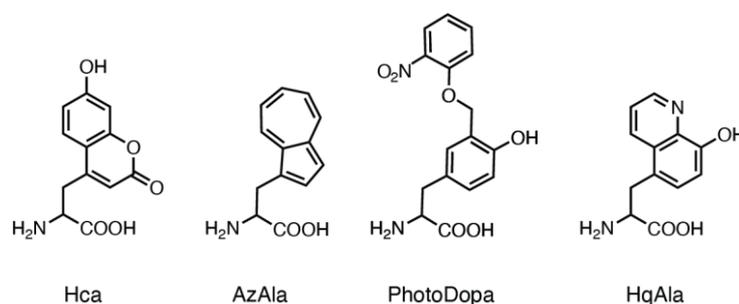


Abbildung 29: Chemische Strukturen der nicht-natürlichen Aminosäuren 7-Hydroxycoumarylalanin (Hca), Azulenalalin (AzAla), *ortho*-Nitrobenzyl Dopamin (PhotoDopa) und 8-Hydroxychinolin Alanin (HqAla).

3.2.2.2 Konstruktion einer auf Substratgruppen fokussierten Zufalls-Bibliothek mittels Sexual PCR

Die Aminoacyl-tRNA Synthetasen HqAlaRS, AzAlaRS und PhotoDopaRS erkennen zu Hca ähnliche Substrate. Die Aminosäure-Sequenzen dieser Aminoacyl-tRNA Synthetasen enthalten damit prinzipiell die Information um sterisch anspruchsvolle nicht-natürliche Aminosäuren mit Doppelringstruktur, so wie Hca, zu erkennen. Zudem erkennt gemäß Literaturangaben keine dieser Aminoacyl-tRNA-Synthetasen, obwohl sie ausgehend von der *Mj*TyrRS engineered wurden, die natürliche Aminosäure Tyr, was einen Vorteil bei deren weiteren Entwicklung bietet. Um diese Information zu nutzen, wurde ausgehend von diesen 3 Aminoacyl-tRNA Synthetasen versucht, in Anlehnung an die evolutionären Prozesse der Natur — Rekombination, Mutation und Selektion — eine Aminoacyl-tRNA Synthetase für die molekulare Erkennung von Hca zu engineeren. Mit Hilfe der von Willem Stemmer entwickelten *in vitro*-Methode der Sexual PCR (Stemmer, 1994), die auf der zufälligen Fragmentierung und Reassemblierung von DNA-Sequenzen mit einer Homologie > 85 % basiert, wurde eine entsprechende Genbibliothek generiert (Abbildung 30; Abschnitt 2.3.3.6).

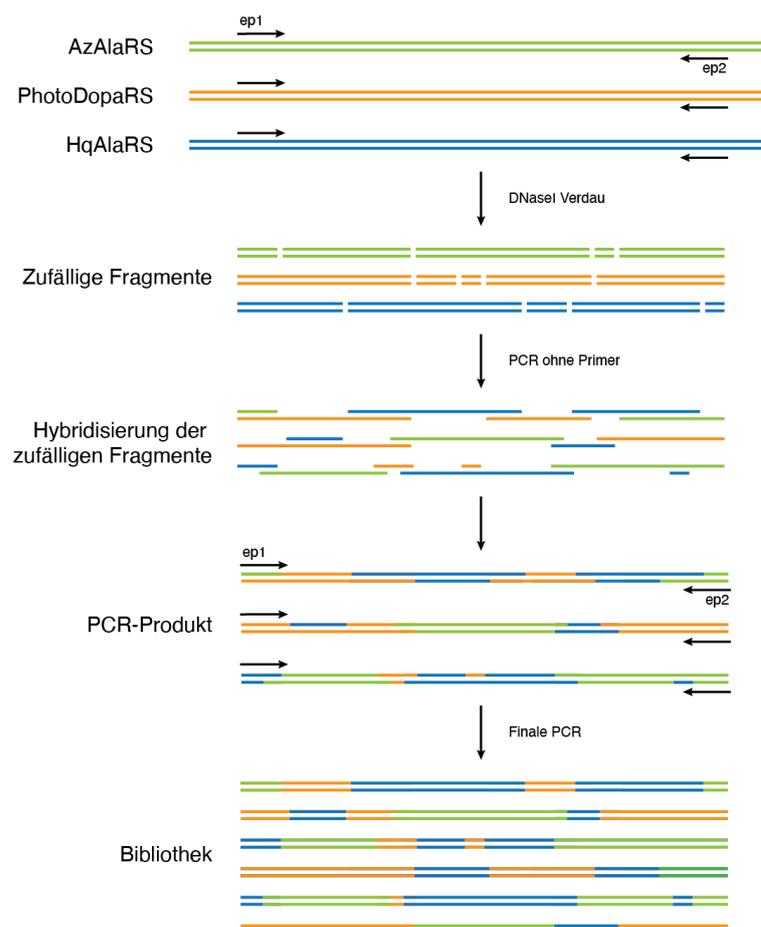


Abbildung 30: Prinzip der sexual PCR (Stemmer, 1994) zur Rekombination von homologen DNA-Sequenzen. Schematisch dargestellt sind die für die sexual PCR verwendeten DNA-Sequenzen von AzAlaRS, PhotoDopaRS und HqAlaRS. Durch den partiellen Verdau mittels DNase I in Gegenwart von Mg^{2+} wurden zufällige Doppelstrangbrüche erzeugt. Die Fragmente werden dann durch eine PCR zu chimären Gensequenzen reassembliert.

Die Aminoacyl-tRNA Synthetasen wurden als doppelsträngige DNA-Fragmente von der Firma Eurofins (Ebersberg, Deutschland) synthetisiert. Dabei deren Planung die Sequenz nicht nur hinsichtlich der für die Expression in *E. coli* günstigen Codons optimiert, sondern auch auf eine möglichst hohe gegenseitige Sequenzhomologie als Grundlage für die effiziente Rekombination geachtet. Für die Generierung der Zufalls-Bibliothek wurde insgesamt 10 µg der für die Aminoacyl-tRNA Synthetasen kodierenden DNA mit 0,7 U DNase I bei 37 °C verdaut. Dabei wurde die Größe der gebildeten DNA-Fragmente über die Inkubationszeit gesteuert. Die Auftrennung der Fragmente erfolgte über ein 2 % Agarosegel, wobei Fragmente mit im Bereich von 35 – 200 bp isoliert wurden. Die Fragmente wurden im Anschluss mit einer PCR ohne *Primer* reassembliert, bei der homologe Fragmente hybridisieren konnten und an den überhängenden Enden jeweils von der Taq DNA-Polymerase zu einem intakten Doppelstrang aufgefüllt wurden. Die anschließende PCR in Gegenwart von flankierenden Primern mit den typischen Schritte aus Denaturierung, Hybridisierung und Elongation führte zu chimären Gensequenzen der ursprünglichen Länge des Gens mit zufällig verteilten Genabschnitten der 3 verwendeten Aminoacyl-tRNA Synthetasen. Als Nebenprodukt erzeugte diese Methode vereinzelt Punktmutationen mit einer Mutationsrate von 0,7/kbp (Stemmer, 1994). Da die Komplexität diese Bibliothek bei vollständiger Randomisierung maximal $7 \cdot 10^4$ Varianten betrug und die Mutationsrate dieser Methode verhältnismäßig gering war, wurde die Bibliothek zur Steigerung der Diversität vor dem ersten Selektionszyklus zusätzlich einer zufälligen Mutagenese mittels *epPCR* (*GeneMorphII – Random Mutagenesis Kit*) unterworfen (Abbildung 31). Das *epPCR*-Produkt wurde mit *EamI* verdaut und anstelle der BpaRS auf den Selektionsvektor pSB16.6c2 kloniert. Nach der Transformation von elektrokompententen *E. coli* MG1655 Zellen mit dem Ligierungsprodukt wurde eine Transformanzahl von $2,7 \cdot 10^9$ Kolonien/µg eingesetzter Plasmid DNA erhalten.

Für die Selektion der Varianten mit dem FACS-Verfahren wurden die zuvor auf Agarplatten kultivierten Transformanten in zwei Flüssigkulturen überführt und bei 30° C und 180 rpm inkubiert. Während der exponentiellen Wachstumsphase ($OD_{550} \sim 0,4$) wurde die Flüssigkultur wie auch eine Kontrollkultur mit 1 mM der nicht-natürlichen Aminosäure (Hca) beziehungsweise mit 1 mM der proteinogenen Aminosäure (Tyr) versetzt. Nach einer Stunde wurde die Expression des Reporterproteins eGFP(39a) durch Zugabe von aTc induziert und nach weiteren 16 h wurde je 1 ml Kultur für die Selektion geerntet. Die eGFP(39a)-Fluoreszenz der mit Hca wie auch der mit Tyr versetzten Kultur wurde mit dem FACS-Gerät detektiert und miteinander verglichen. Basierend auf den beobachteten Unterschieden wurden die zu sortierenden Zellen mittels Anpassung des Sortier-Fensters ausgewählt (Abschnitt 2.6.3). Die sortierten Zellen wurden auf Agarplatten ausplattiert und über Nacht bei 37 °C inkubiert.

	1	10	20	30	40	50	60	70	80		
<i>MjTyrRS</i>	MDEFEm	IKRNTSEI	ISEEELREVL	KKDEKSA	YIGFEP	SGKIHLGH	LQIKKMID	LQNAGFDI	IILLADLH	AYLNQK	GELD
AzalaRS	l.....	G.....	W.....	G.....
PhotoDopaRS	l.....	A.....	I.....	A.....
HqAlaRS	l.....	H.....	V.H.....	G.....
Klon1	l.....	A.....	W.....	G.....
Klon2	l.....	G.....	A.....	I.....	G.....
Klon3	l.....	G.....	W.....	G.....
Klon4	l.....	G.....	W.....	G.....
Klon5	l.....	A.....	I.....	A.....
Klon6	l.....	G.....	I.....	A.....
Klon7	l.....	G.....	T.....	I.....	A.....

	90	100	110	120	130	140	150	160			
<i>MjTyrRS</i>	EIRKIGDYN	KKVFEAMGL	KAKYVYGSE	FQ	LDDKYTLN	VYRLALK	TTLKRARR	SMELIARE	DENPKVAE	VIYPIMQ	VNDIH
AzalaRS	HN.....	A.....
PhotoDopaRS	A.....	AA.....
HqAlaRS	RV.....	N.....
Klon1	RV.....	P.....	AA.....
Klon2	A.....	RV.....	NT.....
Klon3	LA.....	AA.....
Klon4	RV.....	AA.....
Klon5	RV.....	N.....
Klon6	RV.....	R.....	A.....
Klon7	RV.....	N.....

	170	180	190	200	210	220	230	240				
<i>MjTyrRS</i>	YLGVDVA	AVGGMEQR	KIHMLA	RELLPKK	VVICIHN	PVLTGLD	GEGKMSS	SKGNFIA	VDDSP	EIRAKIK	KAYCPAG	VVEGNP
AzalaRS	N.....
PhotoDopaRS	M.....	N.....	Q.....
HqAlaRS	D.G.....
Klon1	M.....	N.....
Klon2	I.G.....
Klon3	M.....	N.....	Q.....
Klon4	M.....	N.....	QH.....	A.....
Klon5	D.G.....	Q.....
Klon6	N.....	Q.....	P.....
Klon7	D.G.....

	250	260	270	280	290	300				
<i>MjTyrRS</i>	IMEIAKYF	LEYPLTI	KRPEKFG	DLTVNSY	EELSLFK	NKELHP	MDLKN	VAEELIK	ILEPIR	KRL
AzalaRS	r.....
PhotoDopaRS	r.....
HqAlaRS	r.....
Klon1	r.....
Klon2	r.....
Klon3	r.....
Klon4	r.....
Klon5	r.....
Klon6	r.....
Klon7	r.....

Abbildung 31: Proteinsequenz-Alignment der für die Bibliothek verwendeten Aminoacyl-tRNA Synthetase-Sequenzen und der mittels Sexual PCR und *ep*PCR generierten Bibliothek im Vergleich zu *MjTyrRS*. Für die AzAlaRS (grün), PhotoDopaRS (orange) und HqAlaRS (blau) spezifische Aminosäuren sind als Ergebnis der DNA-Sequenzierung der Klone 1–7 dargestellt. In schwarzen Großbuchstaben sind die nicht eindeutig zu einer der 3 Aminoacyl-tRNA Synthetasen zuzuordnenden Aminosäuren gekennzeichnet. Mutationen, deren Herkunft auf die *ep*PCR zurückzuführen war, sind in Rot gekennzeichnet. Kleinbuchstaben kennzeichnen einen gezielten Aminosäureaustausch, um ein verkürztes Protein durch ein zweites Start-Methionin (M6L) zu verhindern bzw. um die Interaktion von Aminoacyl-tRNA Synthetase und tRNA zu verbessern (D286R).

Da die Anzahl an Aminoacyl-tRNA Synthetase-Varianten in dieser Bibliothek relativ gering war und jeder Selektions-Zyklus diese weiter reduzierte, wurden die sortierten Aminoacyl-tRNA Synthetase nach dem ersten, zweiten und vierten Zyklus jeweils einer weiteren Zufallsmutagenese unterworfen. Dafür wurden die sortierten Zellen von den Agarplatten geschabt, die Plasmid-DNA isoliert und eine *epPCR* mit den *Primern* ep1 und ep2 mit einer Mutationsrate von 0 – 4 Aminosäureaustauschen pro Aminoacyl-tRNA Synthetase-Sequenz durchgeführt. Im Anschluss daran wurden das *epPCR*-Produkt erneut kloniert, transformiert und wie bereits oben beschrieben kultiviert.

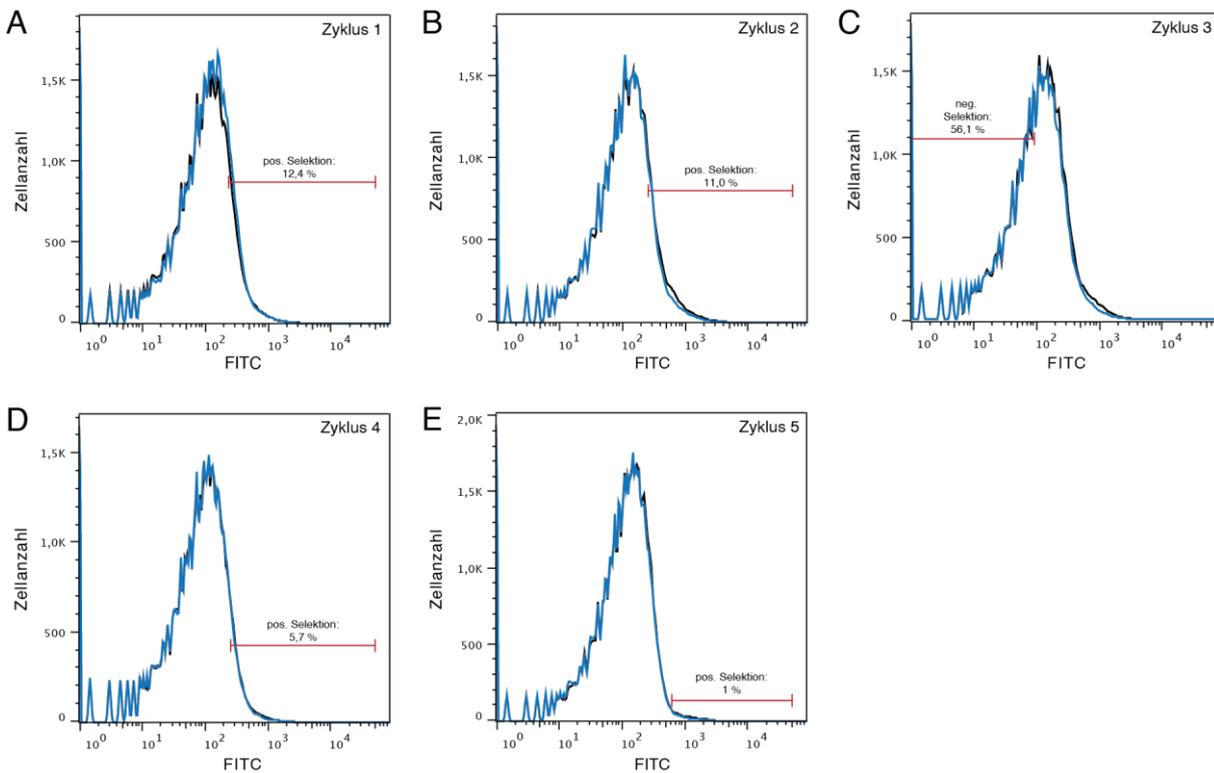


Abbildung 32: Selektionskampagne der Substratgruppen fokussierten Zufalls-Bibliothek. Die aus den Aminoacyl-tRNA Synthetasen AzalaRS, PhotoDopaRS und HqAlaRS durch zufällige Rekombination und Zufallsmutagenese generierte Bibliothek wurde hinsichtlich der Fähigkeit, Hca zu erkennen, selektiert. Dargestellt ist das Fluoreszenzsignal von eGFP(39a) erhalten nach Suppression des Amber-Stop Codons in Gegenwart von 1 mM Hca (blau) bzw. 1 mM Tyr (grau) nach 16-stündiger Expression des Reporterproteins. Der Anteil der sortierten Zellen ist mit einem roten Balken gekennzeichnet und die Prozentangabe bezieht sich auf alle im aktuellen Zyklus analysierten Zellen. Bei einer Positiv-Selektion wurden Zellen aus der Kultur mit Hca und bei der Negativ-Selektion wurden Zellen aus der Kultur mit Tyr selektiert. Für die Messung wurden die Spannung der Laser am FACS-Gerät wie folgt eingestellt: FITC = 550 V; FSC = 300 V; SSC = 700 V

Im Verlauf der Selektion konnte bei dem Vergleich der eGFP(39a)-Fluoreszenz der Kulturen von *E. coli* MG1655 mit den Aminoacyl-tRNA Synthetase-Varianten kein signifikanter Unterschied zwischen Hca und Tyr detektiert werden (Abbildung 32). Lediglich in Zyklus 2 und 3 war eine Veränderung zu beobachten, bei der aber die mit Tyr versetzte Kultur eine geringe Erhöhung der Fluoreszenzintensität (als Rechtsverschiebung der Kurve in Abbildung 32) zeigte. Um Aminoacyl-tRNA Synthetase-Varianten, die Tyr als Substrat erkennen, abzureichern, wurde in Zyklus 3 eine Negativ-Selektion durchgeführt, bei der aus der mit Tyr

versetzen Kultur jene Aminoacyl-tRNA Synthetase-Varianten selektiert wurden, die nur ein geringes Fluoreszenzsignal zeigten. Im Anschluss wurden zwei weitere Positiv-Selektions-Zyklen durchgeführt, wobei beim letzten Zyklus nur 1 % der Zellen mit maximalem Signal selektiert und für die Einzelklonanalyse verwendet wurden.

3.2.2.3 Einzelklonanalyse und Sequenzvergleich der selektierten Aminoacyl-tRNA Synthetasen zur molekularen Erkennung von Hca

Nach dem fünften Zyklus der Selektion (Abbildung 32) des wurden die *E. coli* Zellen mit den Aminoacyl-tRNA Synthetase-Varianten auf Agarplatten ausplattiert.

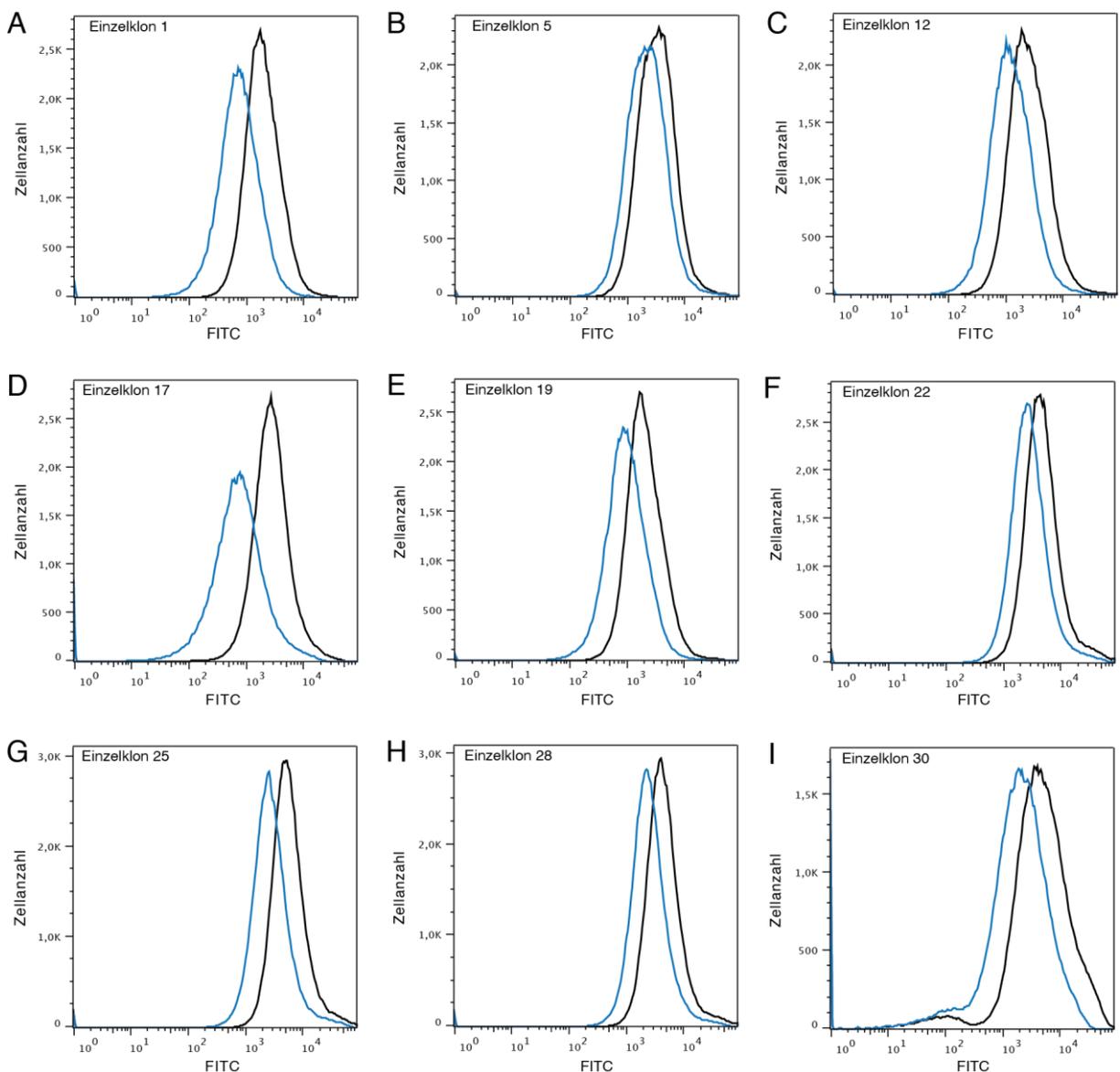


Abbildung 33: Einzelklonanalyse der nach dem fünften Selektions-Zyklus gemäß Abbildung 32 isolierten Klone. Dargestellt ist das Fluoreszenzsignal von eGFP(39a), erzeugt durch die Suppression des Amber-Stop Codons in Gegenwart von 1 mM Hca (blau) und 1 mM Tyr (grau) nach 16-stündiger Expression des Reporterproteins. Für die Messung wurden die Spannungen der Laser am FACS-Gerät wie folgt eingestellt: FITC = 550 V; FSC = 300 V; SSC = 700 V

Für die Einzelklonanalyse zufällig von der Platte gepickte Kolonien wurden separat kultiviert und die Fähigkeit der Aminoacyl-tRNA Synthetase-Variante, Hca zu erkennen, wurde mittels FACS evaluiert (Abschnitt 2.4.1.2 und 2.6.3).

Bei der Einzelklonanalyse wurde die Kultur mit 5 mM Hca oder Tyr versetzt. Nach 16-stündigen Expression des Reporterproteins wurde für jede Kultur die von eGFP(39a) stammende Fluoreszenz quantifiziert. Der Vergleich zwischen der mit Hca und der mit Tyr versetzten Kultur gab Aufschluss über das von der jeweiligen Aminoacyl-tRNA Synthetase-Variante bevorzugte Substrat. Insgesamt wurden 30 Einzelklone analysiert, jedoch wurde von keiner dieser Aminoacyl-tRNA Synthetase-Varianten Hca als Substrat gegenüber Tyr bevorzugt (Abbildung 33).

	1	10	20	30	40	50	60	70	80
<i>MjTyrRS</i>	MDEFEMIKRNTSEIIESEEELREVLKDKDEKSAYIGFEPSPGKIHLGHYLQIKKMIDLQNGAFDIIILLADLHAYLNQKGLD								
AzAlaRS1.....			G			W	G	
PhotoDopaRS1.....			A			I	A	
HqAlaRS1.....			H			V.H	G	
Einzelklon11.....			G			W	A	
Einzelklon51.....			A		M	I	A	
Einzelklon121.....			A			I	G	
Einzelklon301.....			A			I	A	
		90	100	110	120	130	140	150	160
<i>MjTyrRS</i>	EIRKIGDYNKKVFEAMGLKAKYVYGSEFQLDKDYTLNVYRLALKTTLKRARRSMELIAREDENPKVAEVIYPIMQVNDI								
AzAlaRS			HN					A
PhotoDopaRS			A					AA
HqAlaRS			RV					N
Einzelklon1			HN					N
Einzelklon5		D	HN					AA
Einzelklon12		D	HN		M			AA
Einzelklon30			A					N
		170	180	190	200	210	220	230	240
<i>MjTyrRS</i>	YLGVDVAVGGMEQRKIHMIAARELLPKKVVCIHNPVLTGLDGEGKMSSSKGNFIAVDDSPPEIRAKIKKAYCPAGVVEGNP								
AzAlaRS								
PhotoDopaRS	M	N	Q					
HqAlaRS	D.G							
Einzelklon1	D.G	T	R	T				
Einzelklon5	M	N	A					
Einzelklon12	M	N						
Einzelklon30	D.G		E					
		250	260	270	280	290	300		
<i>MjTyrRS</i>	IMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESFLKKNELHPMDLKNVAEELIKILEPIRKRL								
AzAlaRS				r				
PhotoDopaRS				r				
HqAlaRS				r				
Einzelklon1				r				
Einzelklon5				r				
Einzelklon12				r				
Einzelklon30				r				

Abbildung 34: Proteinsequenz-Alignment der sequenzierten Einzelklone im Vergleich zu den für die Bibliothek verwendeten Aminoacyl-tRNA Synthetasen und *MjTyrRS*. Für die AzAlaRS (grün), PhotoDopaRS (orange) und HqAlaRS (blau) spezifische Aminosäuren sind in den Sequenzen der Einzelklone analog koloriert, dargestellt. In einem schwarzen Großbuchstaben ist die nicht eindeutig zu einer der 3 Aminoacyl-tRNA Synthetase zuordenbare Aminosäure gekennzeichnet. Mutationen, deren Herkunft auf die *epPCR* zurückzuführen waren, sind in Rot gekennzeichnet. Kleinbuchstaben kennzeichnen gezielte Aminosäureaustausche, um ein verkürztes Protein durch ein zweites Start-Methionin (M6L) zu verhindern beziehungsweise um die Interaktion von Aminoacyl-tRNA Synthetase und Suppressor-tRNA zu verbessern (D286R).

Eine mögliche Promiskuität der Aminoacyl-tRNA Synthetase-Varianten und die damit verbundene potenzielle molekulare Erkennung von Hca wurde allerdings nicht evaluiert.

Aus vier der 30 Einzelklone wurde die Plasmid-DNA isoliert und deren Aminoacyl-tRNA Synthetase-Sequenzen analysiert (Abbildung 34). Dabei konnten die Aminosäuresequenzen der Einzelklone keinen Aufschluss darüber geben, was auf die bevorzugte Erkennung von Tyr hindeuten würde.

3.3 Unterdrückung der Konkurrenz des Release Faktors 1 um das Amber Stop Codon zur Steigerung der Ausbeute an rekombinantem eGFP(39a) mit einer nicht-natürlichen Aminosäure

Die Suppression des Amber-Stop Codons führt zu deutlich geringerer Ausbeute bei der Expression der nicht-natürlichen Variante des eGFP(39a) im Vergleich zu dem rekombinanten Protein mit durchgängigen Leserahmen. Das wurde bereits bei der Produktion der Protein-Varianten 36Bpa-NFW und 36Y-NFW (Abschnitt 3.1.1) deutlich. Der Grund dafür liegt in der persistenten Konkurrenz des Rf-1 am Ribosom mit der beladenen Suppressor-tRNA um das Amber-Stop Codon (Youngman *et al.*, 2008). Um höhere Ausbeuten an nicht-natürlichen rekombinanten Proteinen zu erhalten, wurde versucht, diese Konkurrenz mittels Genome Engineering zu beseitigen, ohne die Lebensfähigkeit der *E. coli* Zellen zu belasten. Dieses Projekt wurde mit Unterstützung durch Kristina Schira durchgeführt, wobei die Konstruktion der *E. coli* Stammvarianten und die Klonierung des Expressionsplasmides pSB16.6c102 bereits in ihrer Meisterarbeit veröffentlicht wurden (Schira, 2020).

3.3.1 Deletion des Release Faktors 1 in einem *E. coli* B-Stamm

Bei der Termination der Translation erkennt der Rf-1 das am seltensten vorkommende Amber- und das am häufigsten vorkommende Ochre-Stop Codon. Das Ochre-Stop Codon wird zusätzlich vom Rf-2 erkannt, der auch am Opal-Stop Codon die Translation terminiert (Scolnick *et al.*, 1968; Nakamura *et al.*, 1995; Shaw & Green, 2007). Eine Deletion des für den Rf-1 kodierenden *prfA*-Gens führt dazu, dass die Ochre- und das Opal-Stop Codon ausschließlich vom Rf-2 terminiert werden. Die 3 % an Amber-Stop Codons im *E. coli* Genom wird damit nicht mehr erkannt (Abschnitt 1.4).

Die Deletion des *prfA*-Gens in den B-Stamm *E. coli* NEB Express erfolgte durch homologe Rekombination mit Hilfe der Lambda Red Bakteriophagen-Proteine Gam, Beta und Exo, die auf dem Plasmid pKD46 kodiert sind (Datsenko & Wanner, 2000). Eine doppelsträngige Donor-DNA, bestehend aus dem Kanamycin-Resistenz vermittelnden Gen einer Aminoglycosid-Phosphotransferase (*kanR*), welche von einer zum *E. coli* Genom homologen DNA-Sequenz flankiert ist, wurde dabei durch die Exonuklease an beiden 5'-Enden verdaut. Das Lambda Red Bakteriophagen-Protein Beta schützt die durch die Exonuklease erzeugte,

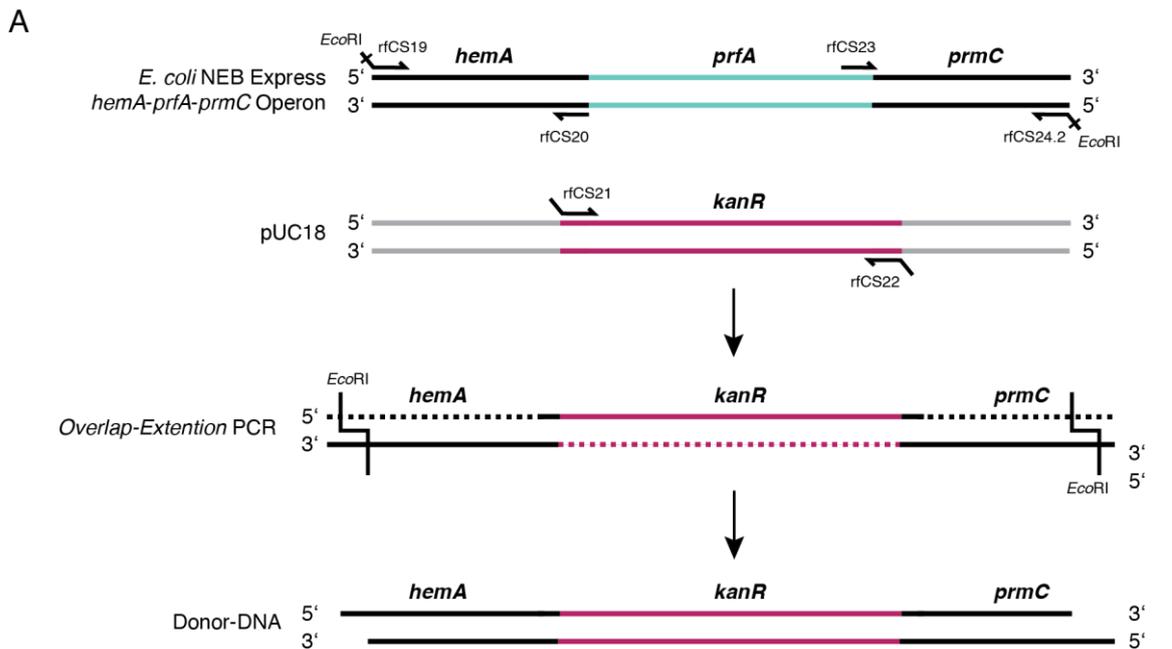
einzelsträngige DNA und fördert die Hybridisierung an die komplementäre, genomische DNA. Das Gam-Protein verhindert den Verdau der einzelsträngigen Donor-DNA durch endogene Exonukleasen (Sharan *et al.*, 2009).

Zur Konstruktion der Donor-DNA wurde das Kanamycin Resistenz vermittelnde Gen, welches *prfA* ersetzen sollte, ausgehend von dem Plasmid 13ABRFYP-1415337 (Vektor der Gensynthese von GeneArt; Erhalten von XL-Protein, Freising) mittels einer PCR mit den *Primern* rfCS21 und rfCS22 amplifiziert. Als homologe Arme, die für die Rekombination essenziell sind, wurden die *prfA* flankierenden Gene *hemA* und *prmC* des *hemA-prfA-prmC*-Operons gewählt (Abbildung 35 A). Dabei ersetzt das Kanamycin Resistenz vermittelnde Gen *prfA* exakt von Start- bis zum Sop-Codon (Abbildung 35 B).

Dafür wurde die genomische DNA aus *E. coli* NEB Express isoliert und als Matrizen-DNA für die Amplifizierung von *hemA* und *prmC* mit den *Primern* rfCS19 und rfCS20 beziehungsweise rfCS23 und rfCS24.2 in getrennten PCR-Ansätzen amplifiziert. Die drei erhaltenen DNA-Fragmente konnten durch die jeweils zueinander homologen endständigen Sequenzen in einer *Overlap-Extension* PCR unter Verwendung der flankierenden *Primer* rfCS19 und rfCS24.2 zu einem DNA-Fragment assembliert und in weiterer Folge amplifiziert werden.

Um die Rekombinationseffizienz zu steigern, wurde nicht die synthetische lineare Donor-DNA verwendet, vielmehr wurde das aus der *Overlap-Extension* PCR erhaltene Produkt über zwei *EcoRI*-Schnittstellen, die sich auf den beiden flankierenden *Primern* befanden, auf dem Plasmid pASK75 kloniert und damit erst *E. coli* NEB Express transformiert. Das ermöglichte auch eine Sequenz-Überprüfung der Donor-DNA mittels Sequenzierung. Zudem führte die *in vivo* Amplifizierung der auf dem Plasmid kodierten Donor-DNA zu geeigneten Methylierung für den gewünschten *E. coli* Stämme. Für die Rekombination wurde das Plasmid pASK75 mit *EcoRI* verdaut, die Donor-DNA mittels einer präparativen Agarose-Gelelektrophorese gereinigt und in elektrokompente *E. coli* NEB Express Zellen, die bereits das Plasmid pKD46 trugen, transfiziert.

Erfolgreich rekombinierte Klone, die also anstelle des *prfA*-Gens das Kanamycin Resistenz vermittelnde Gen trugen, wurden auf Agarplatten mit Kanamycin (35 mg/L) selektiert und im Anschluss durch mehrfaches Ausstreichen auf Agarplatten und Inkubation bei 37 °C von dem Plasmid pKD46 kuriert (Abschnitt 2.2.6). Eine abschließende Überprüfung erfolgte mittels PCR-Amplifizierung der rekombinierten DNA-Sequenz aus der genomischen DNA mit den *Primern* rfCS41 und rfCS42, die außerhalb der Donor-DNA Sequenz hybridisieren, gefolgt von einem analytischen Restriktionsverdau des PCR-Produktes. Der *E. coli* Stamm NEB Express mit deletiertem Rf-1 erhielt die Bezeichnung *E. coli* NEB Express Δ Rf-1.



B

```

NEB Expr.  GAATATTCTGCGCGACAGCCTCGGGCTGGAGTAGcagtacatcattttcttttttttacag
           GluTyrSerAlaArgGlnProArgAlaGlyValAlaValHisHisPheLeuPheLeuGln
           hemA >
ΔRf-1      GAATATTCTGCGCGACAGCCTCGGGCTGGAGTAGcagtacatcattttcttttttttacag
           .....
           hemA >

NEB Expr.  ggtgcatttacgcctATGAAGCCTTCTATCGTTGCCAAACTGGAAGCCCTGCATGAACGC
           GlyAlaPheThrProMetLysProSerIleValAlaLysLeuGluAlaLeuHisGluArg
           < prfA
ΔRf-1      ggtgcatttacgcctATGATTGAACAGGATGGCCTGCATGCGGGTAGCCCGGCAGCGTGG
           .....MetIleGluGlnAspGlyLeuHisAlaGlySerProAlaAlaTrp
           < kanR

//

NEB Expr.  GCCGACCAACTGGCGGCGTTGTCCGAGCAGGAATAATGGAATATCAACACTGGTTACGTG
           AlaAspGlnLeuAlaAlaLeuSerGluGlnGluEnd
           prfA >
           MetGluTyrGlnHisTrpLeuArgG
           < prmC
ΔRf-1      TCGTTTTTATCGTCTGCTGGATGAATTTTTTCTAA-ATGGAATATCAACACTGGTTACGTG
           AlaPheTyrArgLeuLeuAspGluPhePheEnd- .....
           kanR > < prmC

```

Abbildung 35: Schematische Darstellung der Donor-DNA und deren Herstellung. (A) Die genomische DNA von *E. coli* NEB Express diente als Matrizen-DNA für die Amplifizierung der homologen Arme mittels PCR unter Verwendung der Primer rfCS19 und rfCS20 (439 bp lang) beziehungsweise rfCS23 und rfCS24.2 (521 bp lang). Das Kanamycin Resistenz vermittelnde Gen (*kanR*) wurde als Selektionsmarker für erfolgreich Rf-1-deletierte Klone verwendet und wurde ausgehend von 13ABRFYP-1415337 mit den Primern rfCS21 und rfCS22 amplifiziert. Die Assemblierung der drei DNA-Fragmente erfolgte über die jeweils zur Nachbar-Sequenz homologen Abschnitte mittels *Overlap-Extension* PCR. Für die Amplifizierung *in vivo*, wie auch die Sequenzierung wurde die assemblierte Donor-DNA zunächst auf pASK75 kloniert, welche für die Transformation von *E. coli* NEB Express isoliert und mit *EcoRI* verdaut und wurde. (B) DNA-Sequenz von *E. coli* NEB Express und der Donor DNA (ΔRF-1) um das Start- und das Stop-Codon.

3.3.1.1 Cytofluorimetrische Analyse und Vergleich der Expression von eGFP(39a) in den Stämmen *E. coli* NEB Express Δ Rf-1 und *E. coli* NEB Express

Der generierte Stamm *E. coli* NEB Express Δ Rf-1 wurde hinsichtlich der Suppression des Amber-Stop Codons im Vergleich zum ursprünglichen *E. coli* NEB Express Stamm mit Hilfe des Reporterproteins eGFP(39a) (Kuhn *et al.*, 2010) evaluiert. Dafür wurde das Plasmid pSB16.6c2, auf die BpaRS und deren Suppressor-tRNA sowie eGFP(39a) kodiert waren (Edwardraja *et al.*, 2017), in die *E. coli* Zellen transformiert. Für die Expression des Reporterproteins wurde jeweils eine Kolonie der Transformanden gepickt und in Flüssigmedium kultiviert. Die Zugabe von 1 mM D/L-Bpa beziehungsweise von 1 mM-Tyr in den Kontrollkulturen erfolgte bei OD = 0,4, und nach einer Stunde wurde mit aTc die Expression von eGFP(39a) induziert. Die Fluoreszenz des Reporterproteins wurde mit dem FACS-Instrument sowohl nach 3-stündiger als auch nach 16-stündiger Expression quantifiziert, wobei eine höhere Fluoreszenzintensität auf eine quantitativ stärkere Expression von eGFP(39a) und damit auch auf eine verbesserte Suppression des Amber-Stop Codons schließen ließ.

Hierbei führte die Expression von eGFP(39Bpa) in *E. coli* NEB Express Δ Rf-1 zu einem stärkeres Fluoreszenzsignal als in dem kommerziell erhältlichen *E. coli* NEB Express Stamm. Nach 3-stündiger Expression war die Fluoreszenzintensität (FITC Median) um den Faktor $\sim 1,4$ und nach 16-stündiger Expression um den Faktor $\sim 2,2$ gesteigert (Abbildung 36), was direkt auf die Deletion des Rf-1 und somit auf die fehlende Konkurrenz mit der Suppressor-tRNA um das Amber-Stop Codon zurückzuführen war. Aufgrund der Promiskuität der BpaRS konnte nach 16-stündiger Expression von eGFP(39Tyr) mit *E. coli* NEB Express Δ Rf-1 auch ein um den Faktor 1,4 gesteigertes Fluoreszenzsignal im Vergleich zu der Expression in *E. coli* NEB Express detektiert werden (Abbildung 36, unterbrochene Linien).

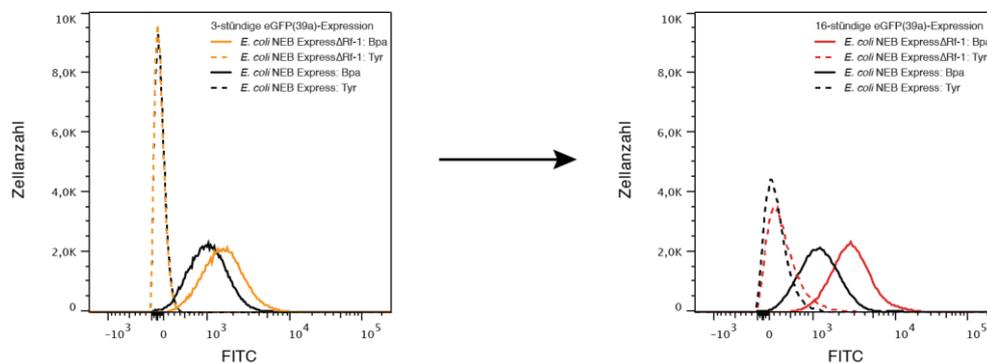


Abbildung 36: Vergleich der *E. coli* Stämme NEB Express Δ Rf-1 und NEB Express. Dargestellt sind die durch die eGFP(39a)-Expression in Gegenwart von Bpa (durchgezogene Linien) und Tyr (unterbrochene Linien) erhaltenen Fluoreszenzsignale nach 3 beziehungsweise 16 Stunden. Ersichtlich ist die Steigerung des Fluoreszenzsignals bei der eGFP(39a)-Expression mit dem *E. coli* Stamm NEB Express Δ Rf-1 (orange/rot) zu NEB Express (schwarz).

3.3.1.2 Produktion von eGFP(39Bpa) in den Stämmen *E. coli* NEB Express Δ Rf-1 und *E. coli* NEB Express, Reinigung und quantitativer Vergleich der Ausbeuten

Abgesehen vom Vergleich der Fluoreszenzintensitäten bei Expression von eGFP(39Bpa) in *E. coli* NEB Express Δ Rf-1 und *E. coli* NEB Express mit Hilfe der BpaRS und in Gegenwart von Bpa wurden die Ausbeuten des gereinigten Reporterproteins verglichen. Die 16-stündige Produktion erfolgte in einer 200-ml LB-Flüssigkultur, der bei $OD_{600} = 0,4$ 1 mM D/L-Bpa zugegeben wurde. Nach einer Stunde wurde mit aTc die Expression von eGFP(39a) induziert. Der Zellaufschluss wurde mit Hilfe eines Ultraschall-Homogenisators durchgeführt und das mit dem *Strep*-tag II fusionierte eGFP(39Bpa) aus dem Zelllysat gereinigt.

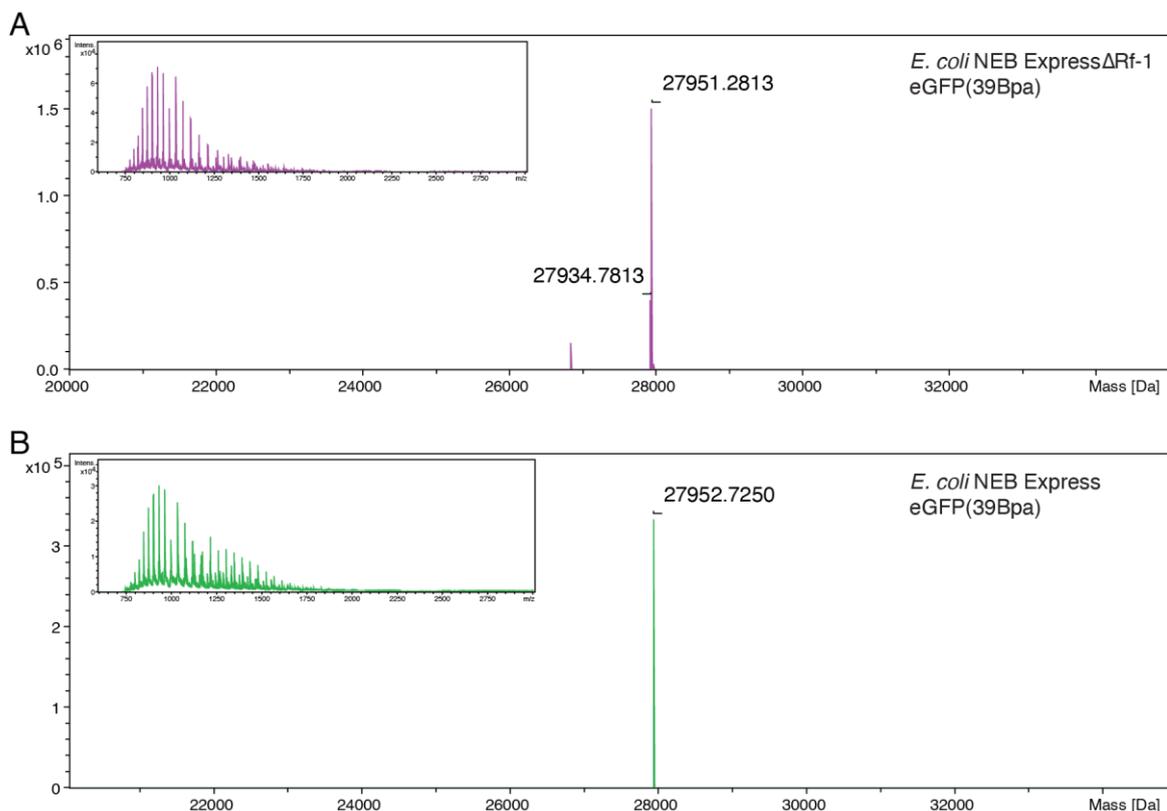


Abbildung 37: ESI-Massenspektrum des gereinigten eGFP(39Bpa). (A) Expression von eGFP(39Bpa) mit *E. coli* NEB Express Δ Rf-1. $MW_{\text{kalk.}} = 27968,5$ Da; $MW_{\text{obs.1}} = 27951,3$ Da; $\Delta m_1 = 17,6$ Da; $MW_{\text{obs.2}} = 27934,8$ Da; $\Delta m_2 = 33,7$ Da. (B) Expression von eGFP(39Bpa) mit *E. coli* NEB Express. $MW_{\text{kalk.}} = 27968,5$ Da; $MW_{\text{obs.}} = 27952,7$ Da; $\Delta m = 15,8$ Da

Die Proteinausbeute wurde spektroskopisch bei 280 nm in Triplikaten ermittelt (Abschnitt 2.5.7) und für eine bessere Vergleichbarkeit auf einen Liter Kultur mit $OD_{550} = 1$ normiert. Die eGFP(39Bpa)-Ausbeute betrug unter Verwendung des *E. coli* Stammes NEB Express Δ Rf-1 $0,35 \pm 0,01$ mg/L \cdot OD und unter Verwendung des *E. coli* Stammes NEB Express $0,15 \pm 0,01$ mg/L \cdot OD. Die um den Faktor 2,3 höhere Ausbeute bei der eGFP(39Bpa)-Expression mit *E. coli* NEB Express Δ Rf-1 entspricht der durch FACS-Analytik ermittelten Steigerung der Fluoreszenzintensität. Die Reinheit des isolierten biosynthetischen

Proteins wurde mittels Massenspektrometrie analysiert und zeigte keine nachweisliche Verunreinigung durch den Einbau von Tyr (Abbildung 37). Die Abweichungen der gemessenen von der berechneten Masse ließen nicht auf den Einbau von Tyr ($\Delta 27,8$ Da) sondern auf einen durch die Bpa-Seitenkette gebildeten Ester ($\Delta 17,6$ Da) beziehungsweise Diester ($\Delta 33,7$ Da) mit benachbarten Aminosäure-Seitenketten schließen (Edwardraja *et al.*, 2017).

3.3.2 Fakultative Inaktivierung des Release Faktors 1 in einem *E. coli* K12-Stamm

Eine Gen-Deletion für Rf-1 (*prfA*), wie sie mit dem *E. coli* B-Stamm NEB Express erfolgreich durchgeführt wurde (Abschnitt 3.1.1), ist mit einem *E. coli* K12-Stamm nicht möglich. Das *prfB*-Gen, das für den Rf-2 kodiert und bei einer Deletion des *prfA*-Gens (Rf-1) allein die Termination an dem am häufigsten vorkommenden Ocher-Stop Codons übernimmt, trägt in *E. coli* K12-Stämmen eine Mutation (A256T), verglichen mit *E. coli* B-Stämmen. Diese Mutation reduziert die Terminationseffizienz des Rf-2 um den Faktor 5 und hat in Kombination mit einem deletierten *prfA*-Gen negative Auswirkungen auf die Zellteilung der *E. coli* K12 Zellen (Johnson *et al.*, 2011; Uno *et al.*, 1996). Deshalb wurde basierend auf der Struktur und der bekannten Funktion der Domänen des Rf-1 (Abschnitt 1.4) der Switch-Loop ausgewählt und diesen dahingehend mutiert, dass er eine TEV-Protease Erkennungssequenz aufwies. So kann die Anzucht der Kultur mit funktionalem Rf-1 erfolgen. Bei der Expression des Ziel-Gens welche simultan mit der Expression der TEV-Protease stattfindet wird der Rf-1 durch den proteolytischen Verdau der TEV-Protease (Kapust *et al.*, 2001) inaktiviert. Diese fakultative Inaktivierung des Rf-1 sollte das bakterielle Zellwachstum nur geringfügig beeinträchtigen und trotzdem die Konkurrenz der RF-1 und der Suppressor-tRNA um das Amber-Stop Codon beenden.

3.3.2.1 Einführung der Erkennungssequenz der TEV-Protease mittels Cas9 und λ -Red Rekombination in den Switch Loop des Release Faktors 1

Für die Mutation des *prfA*-Gens in den *E. coli* K12-Stamm MG1655 zur Einführung der Erkennungssequenz der TEV-Protease (ENLYFQG) wurde der *Switch Loop* des Rf-1 ausgewählt. Der *Switch Loop* ist essenziell für die Funktion des Rf-1 und verbindet die Domäne 2, die das Stop Codon erkennt, mit Domäne 4, gefolgt von Domäne 3, welche die Esterbindung der Peptidyl-tRNA am Ribosom hydrolysiert (Korostelev *et al.*, 2008). Die Genom-Editierung des *prfA*-Gens wurde mit Hilfe der beiden Plasmide pCas (Abbildung 38 (Jiang *et al.*, 2015); Addgene 62225) und pTarget (Abbildung 38 (Jiang *et al.*, 2015); Addgene 62226) und einer gegliederten konstruierten linearen Donor-DNA durchgeführt (Abschnitt 2.2.7). Das Plasmid pCas vermittelt die DNA-Modifizierung durch homologe Rekombination mit der darauf kodierten λ -Red-Bakteriophagen-Proteinen (Gam, Beta, Exo) sowie eine anschließende Selektion gegen das Wildtyp-Genom durch die Exonuklease Cas9 in Kombination mit der auf pTarget kodierten *single-guide* RNA (sgRNA).

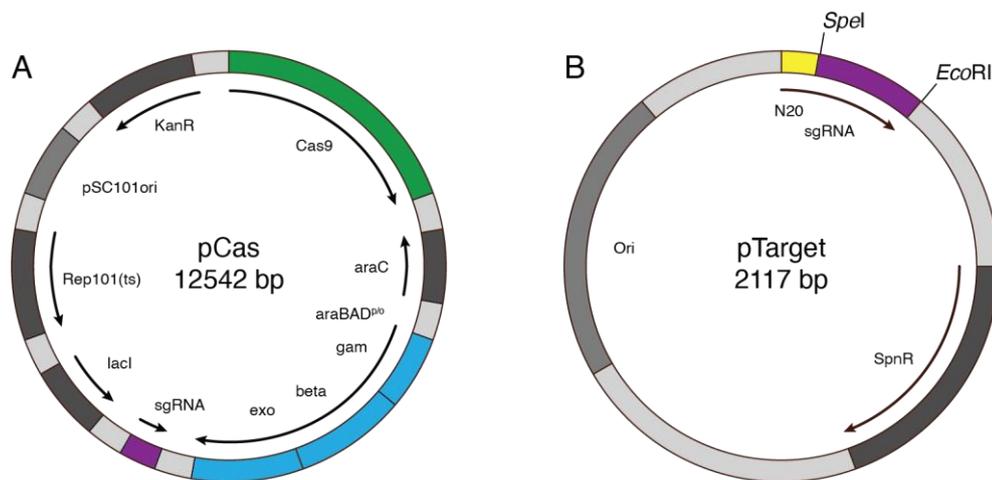


Abbildung 38: Schematische Darstellung der Plasmide pCas und pTarget, die zur Einführung der TEV-Erkennungssequenz in Rf-1 mittels Cas9 und λ -Red Rekombination verwendet wurden. (A) Das Plasmid pCas kodiert für die Endonuklease Cas9 unter der Kontrolle eines konstitutiven Promotors. Zusätzlich trägt das Plasmid die Gensequenz für die λ -Red-Proteine Gam, Beta und Exo unter der Kontrolle des Arabinose induzierbaren Promotor/Operators araBAD. Weitere Elemente auf dem Plasmid sind der pSC101-Replikationsursprung (pSC101ori) und die für die Replikation notwendige Sequenz des temperatursensitiven Rep101-Proteins, Rep101(ts), sowie Gene für den Lac- und Arabinose Repressor (*lacI* bzw. *araC*). Um *E. coli* von pTarget zu kurieren, befindet sich außerdem die Sequenz einer sgRNA mit einem zum Replikationsursprung von pTarget komplementären N20-Protospacer auf dem Plasmid. Das Kanamycin-Resistenzgen ermöglicht die antibiotische Selektion. (B) Das Plasmid pTarget kodiert für die sgRNA samt N20-Protospacer, die komplementär zum Ziellokus der genomischen DNA gewählt wurde. Zusätzlich befindet sich ein Replikationsursprung auf dem Plasmid und das Spectinomycin-Resistenzgen für die antibiotische Selektion.

Die sgRNA (auch bekannt als *guide* RNA, gRNA) besteht aus einer 20 bp umfassenden Sequenz, die homolog zur Zielsequenz im Genom oder Plasmid ist und mit einer PAM-Sequenz (*Protospacer Adjacent Motif*) mit den Basen NGG endet, gefolgt von dem sgRNA-Scaffold, der an Cas9 bindet. Der Komplex aus sgRNA und Cas9 bindet durch Hybridisierung des Protospacers an die genomische DNA und definiert damit die Ziel-Sequenz für den Doppelstrangbruch (Wiedenheft *et al.*, 2012). Ein solcher Doppelstrangbruch in der genomischen oder Plasmid-DNA kann nicht durch endogene Mechanismen der *E. coli* Zelle repariert werden und ist damit lethal.

Die Auswahl der N20 Protospacer-Sequenz erfolgte mit dem webbasierten CRISPR-Tool (<http://crispor.tefor.net>). Dabei wurden die genomische Zielsequenz vorgegeben und passende N20 Protospacer-Sequenzen aus den Vorschlägen ausgewählt (Schira, 2020). In der hier vorliegenden Arbeit ist jedoch nur die erfolgreiche N20 Protospacer-Sequenz 5'-CTACCCGTCGTAACCTGCTG-3' beschrieben. Diese Sequenz hybridisiert mit der zu mutierenden Sequenz im *prfA*-Gen und induziert dort den Doppelsprangbruch. Der Protospacer bindet jedoch nicht an die mutierte *prfA*-Sequenz, was also eine Selektion auf diesen Genotyp ermöglicht. Die Sequenz für den N20-Protospacer und den sgRNA-Scaffold wurde nach Hybridisierung von synthetischen komplementären Oligodesoxynukleotiden auf dem mit *SpeI* und *XbaI* verdauten Vektor pTarget kloniert.

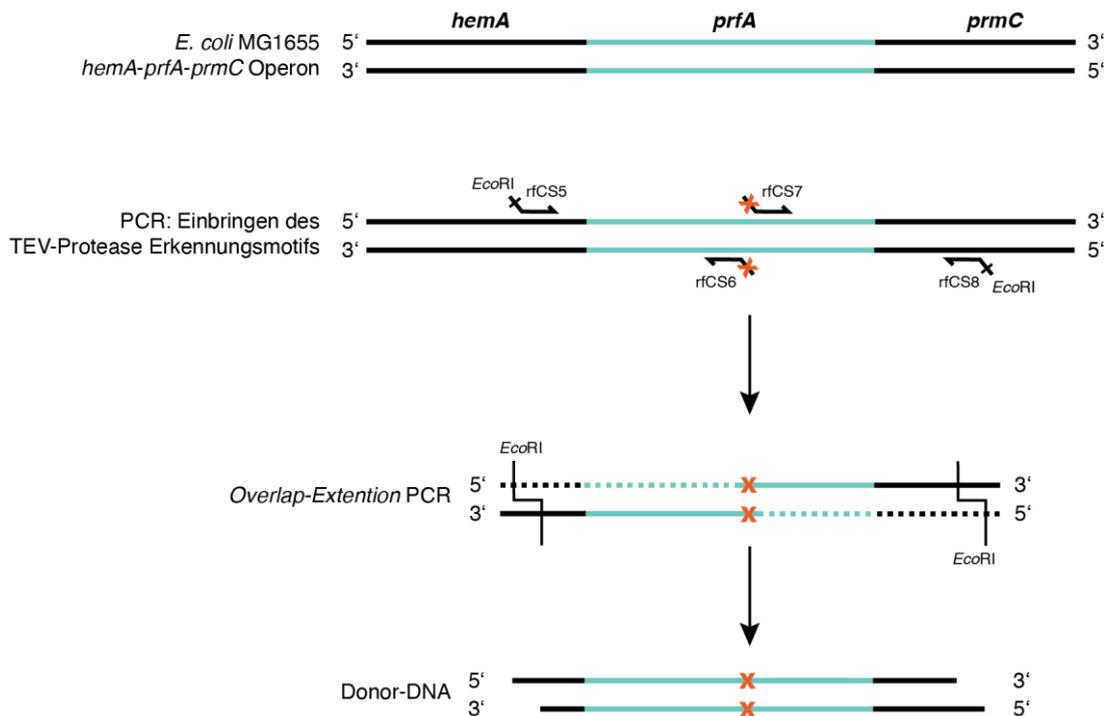


Abbildung 40: Schematische Darstellung der Herstellung der Donor-DNA für die Einführung des TEV-Protease Erkennungsmotifs im Rf-1. Als Matrizen-DNA für die PCR diente die genomische DNA von *E. coli* MG1655. Partiiell komplementäre *Primer*, die zusätzlich für die DNA-Sequenz des TEV-Protease-Erkennungsmotifs kodieren (rotes x), wurden für die Amplifizierung von zwei DNA-Fragmenten verwendet und im Anschluss durch *Overlap-Extension* PCR assembliert. Für die Klonierung wie auch für die Sequenzierung wurde die assemblierte Donor-DNA mit *EcoRI* verdaut und in pUC18 inseriert.

Dabei wurde vor allem die gewünschte Gen-Sequenz, die für das TEV-Protease Erkennungsmotifs kodiert detektiert, jedoch traten in Vorversuchen auch vereinzelt Punktmutationen in dieser Region auf (Abbildung 41).

	870	880	890	900	910	920
Donor-DNA	GAAGCGTCTACCCG	TGAAAATCTCTACTTCCAGGGCGATCGCAGCGACCGTAAC				
	GluAlaSerThrArg	GluAsnLeuTyrPheGlnGlyAspArgSerAspArgAsn				
		TEV Erkennungsmotif				
<i>E. coli</i> MG1655-Tev2a
<i>E. coli</i> MG1655-Tev2b				G
<i>E. coli</i> MG1655-Tev2c				Gly
<i>E. coli</i> MG1655-Tev2d				A
<i>E. coli</i> MG1655-Tev2e				Glu
<i>E. coli</i> MG1655				G
				Arg
				A
				Lys
				CGT . . C . . GCTGGGGAGT
				Arg LeuGlySer

Abbildung 41: Sequenzvergleich der durch die DNA-Editierung generierten Varianten von *E. coli* MG1655 mit der Donor-DNA und der Wildtyp-Sequenz. Das TEV-Protease-Erkennungsmotif in der Donor-DNA ist grün und der korrespondierende Sequenzabschnitt der Varianten ist in grau hinterlegt. Punktmutationen (Abweichungen von der Zielsequenz) und die daraus folgenden Aminosäuremutationen in diesem Sequenzabschnitt sind gekennzeichnet.

Die Ursache für diese Punktmutationen ist nicht bekannt. Eine evolutionär getriebene Mutagenese der *E. coli* Zellen als Reaktion auf die Editierung des an sich essenziellen *prfA*-Gens kann nicht ausgeschlossen werden. Jedoch zeigten die Wachstumskurven der entsprechenden *E. coli* MG1655 Tev-Stammvarianten Tev2a/b/d/e keine signifikante Veränderung im Vergleich zu *E. coli* MG1655 (Schira, 2020).

3.3.2.2 Cytofluorimetrische Analyse und Vergleich der *E. coli* MG1655-Tev Varianten hinsichtlich der Expression von eGFP(39a)

Aufgrund des vergleichbaren Wachstumsverhaltens der *E. coli* MG1655 TEV-Stammvarianten *E. coli* MG1655-Tev2a/b/c (Schira, 2020) wurde die ursprünglich geplante Variante, die keine Mutation innerhalb oder nur eine Punktmutation außerhalb des TEV-Protease-Erkennungsmotifs aufwies, ausgewählt und diese auf die Suppressionsfähigkeit des Amber Stop Codons zu evaluieren. Um wie geplant den Rf-1 an dem eingebrachten TEV-Protease-Erkennungsmotif proteolytisch zu verdauen musste auch eine TEV-Protease exprimiert werden. Dafür wurde das Expressionsplasmid pSB16.6c2 eingesetzt, das im selben Leserahmen für das Reporterprotein eGFP(39a) und die TEV-Protease kodiert und unter der Kontrolle des mit aTc induzierbaren tetA Promotor/Operators steht.

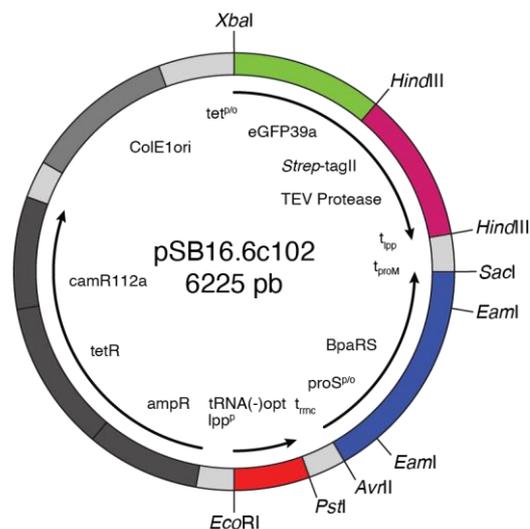


Abbildung 42: Schematische Darstellung des für das Reporterprotein eGFP(39a) und für die TEV-Protease kodierenden Expressionsplasmids. pSB16.6c102 kodiert für eGFP(39a) inklusive des *Strep-tag II* (Skerra & Schmidt, 2000) am 3'-Ende sowie die TEV-Protease unter der Kontrolle des aTc regulierbaren tetA Promotor/Operators (tet^{p^o}) (Skerra, 1994). Zusätzlich trägt das Plasmid die Gensequenz für die BpaRS (Brustad *et al.*, 2008) unter der Kontrolle des konstitutiven proS-Promotors, gefolgt vom proM-Terminator und der dazugehörigen tRNA ($tRNA(-)opt$) unter der Kontrolle des *lpp*-Promotors (lpp^p), in Verbindung mit dem *rnc*-Terminator (t_{rnc}). Die konstitutive Expression der Transkriptionsfusion aus dem Ampicillin-Resistenzgen (AmpR), tet-Repressor (tetR) und dem Chloramphenicol-Resistenzgen — mit einem Amber-Stop Codon an Position 112 — ermöglichte die effektive Repression des tetA Promotor/Operators (tet^{p^o}) für die Expression von eGFP(39a) wie auch eine generelle und eine suppressionsabhängige antibiotische Selektion auf Chloramphenicol.

Um diese Klonierung mittels Restriktionsendonukleasen zu vereinfachen wurde zuvor die *HindIII*-Schnittstelle, die sich auf dem pRK793 Plasmid (Kapust *et al.*, 2001) in der für die

TEV-Protease kodierenden Sequenz befand, mittels einer *Quick-Change* Mutagenese (Abschnitt 2.3.3.3) mit den *Primern* rfCS30 und rfCS31 entfernt. Im Anschluss wurde diese Sequenz durch PCR mit den *Primern* rfCS37 und rfCS38 amplifiziert. Dabei wurde durch den *Primer* rfCS37 eine *HindIII*-Schnittstelle, die *OmpA*-RBS und das Start-ATG am 5'-Ende und durch den *Primer* rfCS38 die DNA-Sequenz für eine weitere *HindIII*-Schnittstelle am 3'-Ende an das PCR-Produkt angefügt. Dieses PCR-Produkt wurde mit *HindIII* verdaut und mit dem ebenfalls mit *HindIII* verdauten und zusätzlich dephosphorylierten Vektor pSB16.6c2 ligiert (Abschnitt 2.3.2 und 2.3.4). Das daraus resultierende Plasmid pSB16.6c102 (Abbildung 42) wurde durch DNA-Sequenzierung verifiziert und in die *E. coli* Stämme MG1655-Tev2a, MG1655-Tev2b und MG1655-Tev2c sowie *E. coli* MG1655 transfiziert. Jeweils ein Klon wurde in einer Flüssigkultur kultiviert. Bei $OD_{550} = 0,4$ wurden die Kulturen mit 1 mM D/L-Bpa versetzt und nach einer Stunde erfolgte die Induktion der Expression von eGFP(39Bpa) wie auch der TEV-Protease mit aTc.

Die Auswirkung der am Rf-1 vorgenommenen Editierung beziehungsweise der aufgetretenen Punktmutationen in Kombination mit der proteolytischen Inaktivierung des Rf-1 durch die in den Zellen produzierte TEV-Protease auf die Amber-Stop Codon Suppression wurde mit Hilfe des Reporterproteins eGFP(39Bpa) evaluiert (Abbildung 43 B). Dabei wurde die durch die eGFP(39Bpa)-Expression in den *E. coli* Stämmen MG1655-Tev2a, MG1655-Tev2b, MG1655-Tev2c wie auch in MG1655 vermittelte Fluoreszenz mit dem FACS-Gerät detektiert und verglichen. Unterschiede in der Fluoreszenzintensität waren auf die in den Zellen vorliegende Anzahl an eGFP(39Bpa) Molekülen zurückzuführen und gaben somit Aufschluss über die Suppression des Amber-Stop Codons in den jeweiligen *E. coli* Stämmen. Die eGFP(39Bpa)-Fluoreszenzsignale, erhalten nach 3-stündiger Expression in den *E. coli* Stämmen MG1655-Tev2a, MG1655-Tev2b und MG1655, zeigten keine signifikanten Unterschiede. Ein um den Faktor 1,8 geringeres Signal wurde dagegen mit in *E. coli* MG1655-Tev2c detektiert (Abbildung 43 B; links). Die Ko-Expression des Reporterproteins und der TEV-Protease über Nacht veränderte dieses Ergebnis jedoch. Hierbei führten die *E. coli* Stämme MG1655-Tev2a, MG1655-Tev2b und MG1655-Tev2c zu erhöhter Expression von eGFP(39Bpa), ersichtlich anhand der Steigerung der Fluoreszenzintensität um den Faktor ~ 7 , im Vergleich zu *E. coli* MG1655, dessen Fluoreszenzsignal stagnierte und im Rahmen der nach 3-stündiger eGFP(39Bpa)-Expression blieb (Abbildung 43 B; rechts).

Des Weiteren wurde die Auswirkung der Genom-Editierung am Rf-1 ohne die durch die TEV-Protease herbeigeführte Proteolyse des Rf-1 untersucht. Dafür wurde eGFP(39Bpa) mit dem Plasmid pSB16.6c2 in den *E. coli* Stämmen MG1655-Tev2a, MG1655-Tev2b, MG1655-Tev2c wie auch in MG1655 exprimiert und einer FACS-Analyse unterzogen (Abbildung 43, A). Dabei stellte sich heraus, dass bereits die Genom-Editierung am Rf-1 alleine nach 3-stündiger eGFP(39Bpa)-Expression die Suppression des Amber-Stop Codons um den Faktor $\sim 2,5$ steigert. Zu erkennen war dies durch die Rechtsverschiebung der Fluoreszenzsignale, welche

durch die eGFP(39Bpa)-Expression in den *E. coli* Stämmen MG1655-Tev2a, MG1655-Tev2b und MG1655-Tev2c erzeugt wurden (Abbildung 43 A: links). Da dies jedoch bei den beschriebenen 3-stündigen eGFP(39Bpa)-Expressionen in den verwendeten Stämme in Kombination mit der TEV-Protease nicht der Fall war, war zu vermuten, dass sich die Anwesenheit der TEV-Protease vor allem zu Beginn negativ auf die Expression des Reporterproteins auswirkte. Diese Hypothese wurde von den Fluoreszenzsignalen, die aus der 16-stündigen eGFP(39Bpa)-Expression bei abwesender TEV-Protease resultierten, gestützt, da diese um den Faktor $\sim 2,5$ niedriger waren als die Fluoreszenzsignale nach 16 Stunden in Gegenwart der TEV-Protease (Abbildung 43, Vergleich der beiden Histogramme rechts).

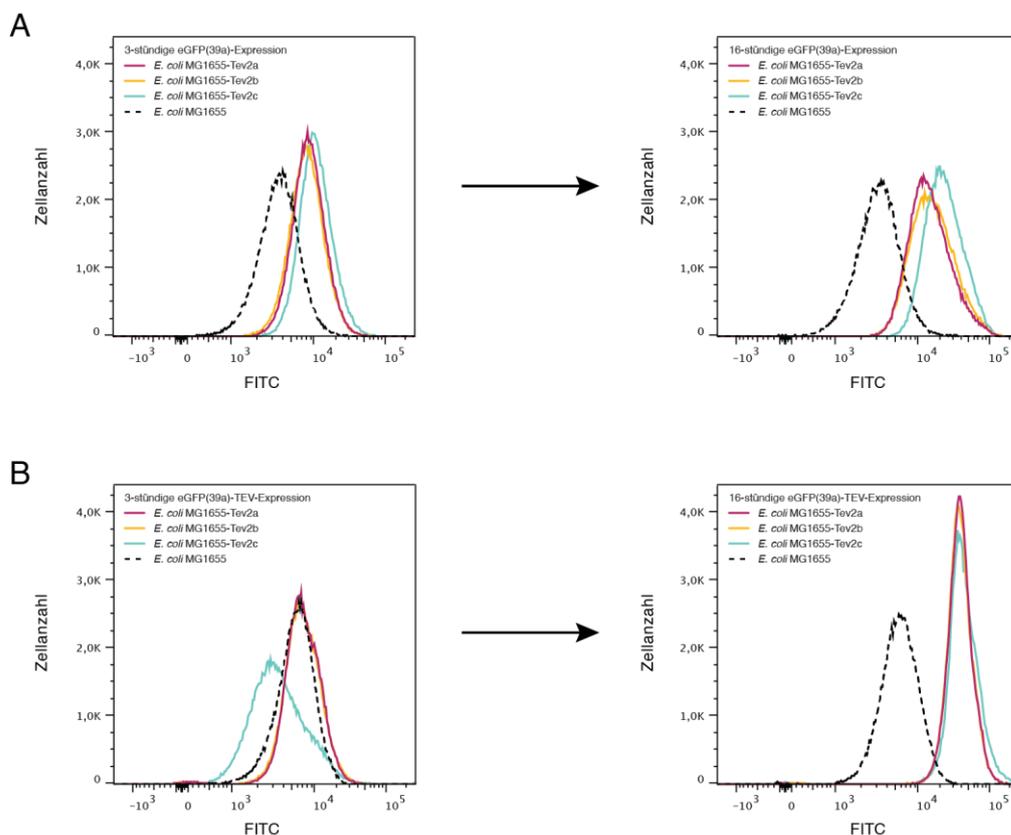


Abbildung 43: Vergleich der *E. coli* Stamm-Varianten MG1655-Tev2a, MG1655-Tev2b, MG1655-Tev2c und MG1655 hinsichtlich der Suppression des Amber-Stop Codons mit Hilfe des Reporterproteins eGFP(39a). Dargestellt sind die durch die Expression von eGFP(39Bpa) vermittelten und mit dem FACS-Gerät detektierten Fluoreszenzsignale. (A) Expression des Reporterproteins mit dem Plasmid pSB16.6c2, (B) Ko-Expression des Reporterproteins und der TEV-Protease mit dem Plasmid pSB16.6c102.

Zu erwähnen ist auch der Einfluss des Plasmides auf die Expression des Reporterproteins und folglich auf die Fluoreszenzintensität (Abbildung 43, Vergleich der schwarzen, unterbrochenen Linien in A und B). Ein Vergleich der *E. coli* MG1655 Kulturen, die entweder das Plasmid pSB16.6c2 oder pSB16.6c102 trugen, zeigte eine um den Faktor 1,8 gesteigerte eGFP(39Bpa)-Fluoreszenz bei der Expression mit pSB16.6c102, das nur zusätzlich die DNA-Sequenz für die TEV-Protease trug. Das bedeutet aber auch, dass die Steigerung der Fluoreszenz nicht

ausschließlich auf die Aktivität der TEV-Protease zurückzuführen war. Es könnte auch die Kombination aus der durch die Genom-Editierung am Rf-1 verringerte Terminationseffizienz mit der damit verbundenen Steigerung der Suppression am Amber-Stop Codon und der noch ungeklärte Effekt des Plasmides pSB16.6c102 dafür verantwortlich sein.

In den folgenden Experimenten wurde mit *E. coli* MG1655-TeV2a weiter gearbeitet, da die 3-stündige eGFP(39Bpa)-Expression in diesem Stamm ein stärkeres Fluoreszenzsignal zeigte als *E. coli* MG1655-TeV2c. Das Fluoreszenzsignal, aufgrund der eGFP(39Bpa)-Expression in *E. coli* MG1655-TeV2a und MG1655-TeV2b erwies sich als gleich stark.

3.3.2.3 Cytofluorimetrische Analyse und Vergleich der Expression von eGFP(39a) mit den Stämmen *E. coli* MG1655-TeV2a und *E. coli* MG1655

Um sicherzustellen, dass die Suppression des Amber-Stop Codons in *E. coli* MG1655-TeV2a nicht nur in Kombination mit BpaRS/Bpa zu einer höheren eGFP(39Bpa)-Fluoreszenzintensität und der zugrundeliegenden Steigerung der eGFP(39Bpa)-Expression führt, wurde die Suppression des Amber-Stop Codons auch in Kombination mit den Synthetasen/Aminosäuren ApaRS (Reichert *et al.*, 2015) und p-Aethylphenylalanin (Apa) sowie MbPylRS (Srinivasan *et al.*, 2002) und Tetrahydrofuran-Lysin (THF-K) evaluiert. Dabei wurde wie in Abschnitt 3.3.2.2 die Suppression des Amber-Stop Codons mit dem modifizierten Rf-1 in Gegenwart und in Abwesenheit der TEV-Protease untersucht. Um ein Expressionsplasmid zu erhalten, womit Apa in eGFP(39a) eingebaut werden konnte, wurde das Gen der BpaRS auf dem für eGFP(39a) kodierenden Plasmid pSB16.6c2 und dem für die eGFP(39a) mit der TEV-Protease kodierenden Plasmid pSB16.6c102 (Abbildung 42) jeweils gegen die ApaRS, welche von dem Plasmid pSB15.23c2 stammte, ausgetauscht. Dies erfolgte mittels Restriktionsklonierung mit den Restriktionsendonukleasen *AvrII* und *SacI*. Die daraus resultierenden Plasmide trugen die Namen pSB16.23c2 und pSB16.23c102. Ein Austausch der zur Aminoacyl-tRNA Synthetase korrespondierenden Suppressor-tRNA war in diesem Fall nicht notwendig, da die ApaRS wie die BpaRS von der *MjTyrRS* abstammte und beide die gleiche tRNA als Substrat erkennen.

Für die Klonierung des Expressionsplasmides, das es ermöglicht, THF-K in eGFP(39a) einzubauen, mussten sowohl die Aminoacyl-tRNA Synthetase als auch die Suppressor-tRNA ausgetauscht werden. Dabei wurden zunächst die PylRS, ebenfalls mittels Restriktionsklonierung unter Zuhilfenahme der Enzyme *AvrII* und *SacI*, anstelle der BpaRS auf die Plasmide pSB16.6c2 und pSB16.6c102 kloniert. In einem zweiten Schritt wurde die für die PylRS spezifische Suppressor-tRNA durch Restriktionsklonierung mit den Enzymen *PstI* und *EcoRI* gegen die für die *MjTyrRS* spezifische tRNA ausgetauscht. Die daraus resultierenden Plasmide trugen in Bezug auf die darauf kodierte Aminoacyl-tRNA Synthetase und die Suppressor-tRNA die Namen pSB16.100d2 und pSB16.100d102.

Für die Untersuchung des *E. coli* Stammes MG1655-Tev2a hinsichtlich der Suppression des Amber-Stop Codons unter Verwendung der Aminoacyl-tRNA Synthetasen BpaRS, ApaRS und PylRS wurden die Plasmide pSB16.6c2, pSB16.6c102, pSB16.23c2, pSB16.23c102, pSB16.100d2 oder pSB16.100d102 in *E. coli* MG1655-Tev2a und *E. coli* MG1655 transfiziert. Jeweils ein Einzelklon wurde für die Anzucht einer bakteriellen Kultur gepickt, welche bei einer $OD_{550} = 0,4$ mit 1 mM der entsprechenden nicht-natürlichen Aminosäure versetzt wurde. Um zu überprüfen, ob der Fehleinbau von natürlichen Aminosäuren bei verminderter Konkurrenz von Rf-1 um das Amber-Stop Codon gesteigert ebenfalls wird, wurde als Kontrolle von jeder mit einer nicht-natürlichen Aminosäure versetzten Kultur eine identische Kultur in Gegenwart einer natürlichen Aminosäure kultiviert. Dabei wurde für die BpaRS und ApaRS exprimierenden Kulturen 1 mM Tyr und für die PylRS exprimierende Kultur 1 mM Lys verwendet. Eine Stunde nach Zugabe der nicht-natürlichen Aminosäuren oder der natürlichen Aminosäuren wurde die Expression des Reporterproteins eGFP(39a) beziehungsweise der TEV-Protease mit aTc induziert. Nach 3-stündiger beziehungsweise 16-stündiger Expression des Reporterproteins wurden die Fluoreszenzsignale mit dem FACS-Gerät detektiert (Abschnitt 2.6.3).

Ein Vergleich der Fluoreszenzintensitäten, welche mit der Expression von eGFP(39a) in *E. coli* MG1655 beziehungsweise der Expression von eGFP(39a) in Kombination mit der TEV-Protease in *E. coli* MG1655-Tev2a erzeugt wurden, gab Aufschluss inwieweit sich die Suppression des Amber-Stop Codons mit der Strategie der fakultativen Inaktivierung des Rf-1 steigern ließ. Dabei fiel auf, dass dadurch die Expression des eGFP(39a) unter Verwendung verschiedener Aminoacyl-tRNA Synthetase zwar gesteigert werden konnte, jedoch nicht im gleichen Maße (Abbildung 44). Der stärkste Unterschied konnte in Kombination mit BpaRS/Bpa nachgewiesen werden. Hier war das Fluoreszenzsignal bei Expression von eGFP(39Bpa)-TEV-Protease in *E. coli* MG1655-Tev2a nach 3-stündiger Expression um den Faktor 2,2 und nach 16-stündiger Expression um den Faktor 14,1 gesteigert (Abbildung 44 A), verglichen mit der eGFP(39Bpa) Expression in *E. coli* MG1655.

Ein Experiment, bei dem eGFP(39Bpa) in Abwesenheit der TEV-Protease in *E. coli* MG1655-Tev2a exprimiert wurde, zeigte eine Steigerung der Fluoreszenzintensität um den Faktor 2,2 nach 3-stündiger und den Faktor 4,6 nach 16-stündiger Expression (Abbildung 44 D), verglichen mit der eGFP(39Bpa) Expression in *E. coli* MG1655. Daraus wurde geschlossen, dass die Proteolyse des RF-1 durch die TEV-Protease in den ersten 3 Stunden der Expression noch keinen nennenswerten Einfluss auf die Suppression des Amber-Stop Codons hatte. Die Steigerung des Fluoreszenzsignals beruhte also vermutlich auf der vorgenommenen Mutation am RF-1, welche die Terminationseffizienz negativ beeinflusst und dadurch die Konkurrenz um das Amber-Stop Codon generell verringert.

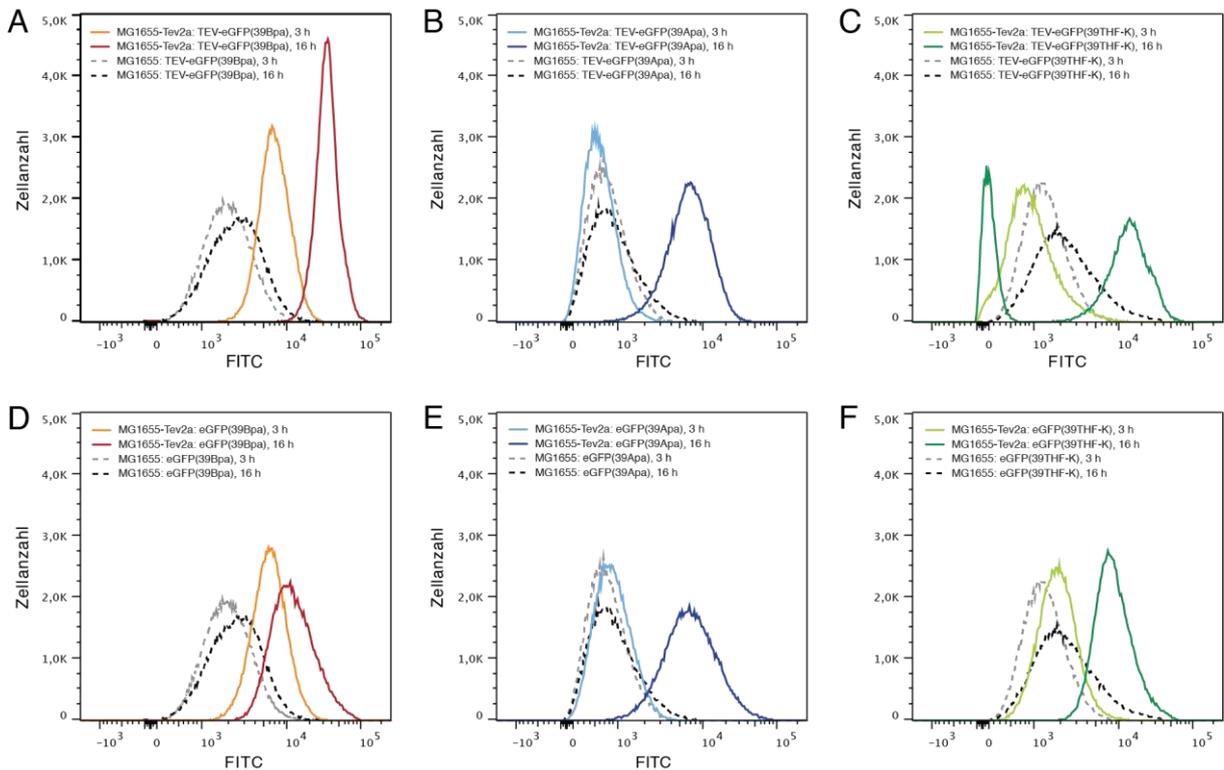


Abbildung 44: Vergleich der eGFP(39a)-Fluoreszenz, nach Amber-Stop Codon Suppression in den *E. coli* Stämmen MG1655-Tev2a und MG1655. Gegenüberstellung der beiden Expressionsstrategien — biclonale Expression von eGFP(39a) mit der TEV-Protease in *E. coli* MG1655-Tev2a und eGFP(39a)-Expression in *E. coli* MG1655 — in Kombination mit den Aminoacyl-tRNA Synthetasen BpaRS (A), ApaRS (B) und PylRS (C) und deren nicht-natürlichen Aminosäure-Substrate.

Analoge Versuche wurden auch mit den Aminoacyl-tRNA Synthetasen ApaRS in Kombination mit Apa und PylRS in Kombination mit THF-K durchgeführt. Dabei wurde deutlich, dass die Steigerung der Fluoreszenz nicht nur von der Konkurrenz um das Amber-Stop Codon abhängt, sondern auch die Aminoacyl-tRNA Synthetase und deren jeweiligen Aminosäure-Substrat eine Rolle spielen. Bei der Expression von eGFP(39Apa) mit der TEV-Protease in *E. coli* MG1655-Tev2a konnte das Fluoreszenzsignal erst nach 16-stündiger Expression um den Faktor 7,6 gesteigert werden (Abbildung 44 B), verglichen mit der eGFP(39Apa) Expression in *E. coli* MG1655. Das entsprach der Steigerung der Fluoreszenzintensität, die nach 16-stündiger eGFP(39Apa)-Expression in *E. coli* MG1655-Tev2a beobachtet wurde (Abbildung 44 E). Bei der 3-stündigen Expression von eGFP(39Apa) mit der TEV-Protease in *E. coli* MG1655-Tev2a wurde sogar eine etwas geringere Fluoreszenz detektiert im Vergleich der Expression in *E. coli* MG1655 wie auch zur eGFP(39Apa)-Expression in *E. coli* MG1655-Tev2a (Abbildung 44 B). Ähnliches wurde bei Verwendung der PylRS in Kombination mit THF-K beobachtet. Die Steigerung der Fluoreszenzintensität, gemessen nach 16-stündiger Expression von eGFP(39THF-K)-TEV-Protease in *E. coli* MG1655-Tev2a wie auch die eGFP(39THF-K)-Expression im gleichen Stamm war im Vergleich zu der eGFP(39THF-K)-Expression in *E. coli* MG1655 um den Faktor 4,3 gesteigert (Abbildung 44 C und F). Das bedeutet, dass die

Expression der TEV-Protease und die damit ausgelöste Proteolyse des Rf-1 keine Auswirkung auf die Steigerung der Expression des Reporterproteins hatte, sondern diese auf die Mutationen im Rf-1 zurückzuführen waren.

Das in Abbildung 44 (C) dargestellte Fluoreszenzsignal der MG1655-Tev2a Kultur, welche eGFP(39a) mit der TEV-Protease exprimiert (dunkelgrün), zeigt, dass einige Zellen nur sehr gering fluoreszierten. Das weiß auf eine separate Zellpopulation mit vermutlicher Beeinträchtigung der Expression beziehungsweise Lebensfähigkeit. Allerdings wurde das in weiterer Folge nicht noch einmal in dieser Intensität beobachtet.

Zusätzlich wurden identische Kontrollkulturen, die mit einer zu der jeweiligen Aminoacyl-tRNA Synthetase korrespondierenden natürlichen Aminosäure versetzt wurden, verglichen. Hierbei wurde deutlich, dass die verminderte Konkurrenz um das Amber-Stop Codon in *E. coli* MG1655-Tev2a nicht ausschließlich den Einbau der nicht-natürlichen Aminosäure fördert, sondern im Falle einer promiskuitiven Aminoacyl-tRNA Synthetase natürlich auch den Fehleinbau durch eine natürliche Aminosäure. Zu erkennen ist dies durch die Verschiebung des Fluoreszenzsignals bei Verwendung des *E. coli* Stammes MG1655-Tev2a im Vergleich zu *E. coli* MG1655 (Abbildung 45, schwarze Linie und schwarze unterbrochene Linie). Eine zusätzliche Expression der TEV-Protease steigert das erzeugte Fluoreszenzsignals (Abbildung 45, rot, blau grün). Am prominentesten ist dies bei der Expression von eGFP(39Bpa) mit der TEV-Protease zu erkennen (Abbildung 45 A), was auf der Promiskuität der BpaRS zurückzuführen ist.

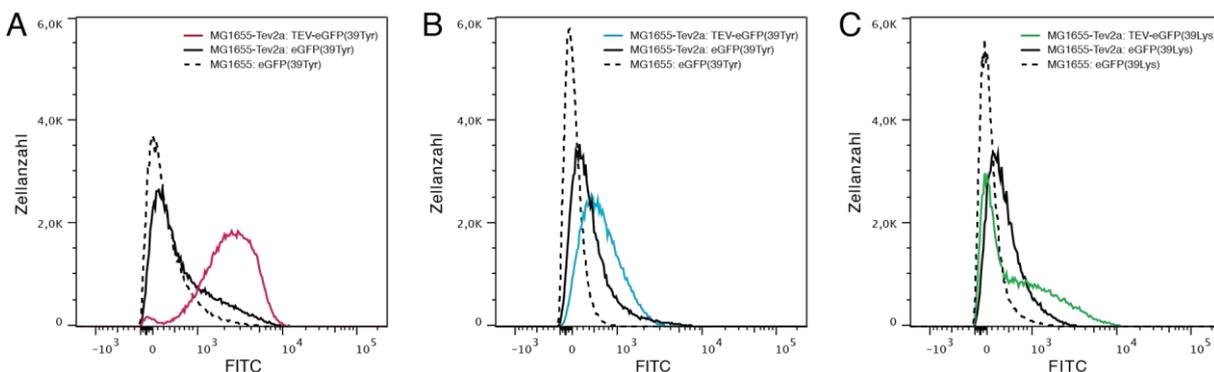


Abbildung 45: Vergleich der eGFP(39a)-Fluoreszenz, gemessen in Anwesenheit der zu den Aminoacyl-tRNA Synthetasen korrespondierenden natürlichen Aminosäuren. Exprimiert wurde eGFP(39a) in *E. coli* MG1655 oder *E. coli* MG1655-Tev2a sowie eGFP(39a) mit der TEV-Protease in *E. coli* MG1655-Tev2a. Beide Kulturen, die ein Plasmid trugen, welches für die ausgehend von *Mj*TyrRS engierte Aminoacyl-tRNA Synthetasen BpaRS (A) oder ApaRS (B) kodierte, wurden mit Tyr versetzt. Als natürliche Aminosäure wurde der Kultur, welche aus *E. coli* Zellen mit dem für die PylRS kodierenden Plasmid besteht (C), Lys zugesetzt.

3.3.2.4 Expression von eGFP(39a) mit den Stämmen *E. coli* MG1655-Tev2a und *E. coli* MG1655, Reinigung und quantitativer Vergleich der Ausbeuten

Der Vergleich der durch die Expression des eGFP(39a)-Reporterproteins mit *E. coli* MG1655-Tev2a und *E. coli* MG1655 erhaltene Fluoreszenzsignale gab zwar Aufschluss über die relative Steigerung der Expression durch die effizientere Suppression des Amber-Stop Codons, aber es konnte weder eine Aussage über die absolute Ausbeutensteigerung von eGFP(39a) getroffen werden, noch über die genaue Zusammensetzung des Proteins.

Um die Expression von eGFP(39a) zusammen mit der TEV-Protease in *E. coli* MG1655-Tev2a und eGFP(39a) in *E. coli* MG1655 vergleichen zu können wurde das Reporterprotein produziert und gereinigt. Dafür wurden die mit pSB16.6c2, pSB16.23c2 oder pSB16.100d2 transfizierten *E. coli* MG1655 Zellen und die mit pSB16.6c102, pSB16.23c102 oder pSB16.100d102 transfizierten *E. coli* MG1655-Tev2a Zellen in Gegenwart der jeweiligen nicht-natürlichen Aminosäure in 200 mL Flüssigmedium kultiviert. Die Zugabe der nicht-natürlichen Aminosäuren erfolgte bei $OD_{550} = 0,4$ und eine Stunde danach wurde die Expression von eGFP(39a) beziehungsweise eGFP(39a) mit der TEV-Protease für 16 Stunden mit aTc induziert. Nach dem Zellaufschluss mit einem Ultraschall-Homogenisator wurde das mit dem *Strep*-tag II fusionierte eGFP(39a) durch Streptavidin-Affinitätschromatographie aus dem Zellysate gereinigt. Die Proteinausbeute wurde spektroskopisch, bei 280 nm mit dem Nanodrop-Gerät, in Triplikaten ermittelt und zwecks Vergleichbarkeit auf einen Liter Kultur und $OD_{550} = 1$ normalisiert.

Aus der *E. coli* MG1655 Kultur mit dem Plasmid pSB16.6c2 wurde $0,33 \pm 0,02$ mg/L•OD eGFP(39Bpa) und aus der *E. coli* MG1655-Tev2a Kultur mit dem Plasmid pSB16.6c102 $1,43 \pm 0,003$ mg/L•OD eGFP(39Bpa) gereinigt. Das entsprach einer 4,4-fachen Steigerung der Ausbeute bei Verwendung von *E. coli* MG1655-Tev2a in Kombination mit dem für eGFP(39a) und die TEV-Protease kodierenden Plasmid. Die Reinheit der Probe wurde mittels ESI-Massenspektrometrie überprüft und zeigte keine nachweisbaren Verunreinigungen durch den Einbau von Tyr. Die Abweichungen der gemessenen zu der berechneten Masse nicht auf einen Einbau von Tyr ($\Delta 27,8$ Da), sondern auf einen durch die Bpa-Seitenkette formierten Ester ($\Delta 17,0$ Da) beziehungsweise Diester ($\Delta 35,1$ Da) mit einer benachbarten Aminosäure-Seitenkette schließen (Abbildung 46 A und B).

Aus der *E. coli* MG1655 Kultur mit dem Plasmid pSB23.6c2 wurde $0,18 \pm 0,004$ mg/L•OD eGFP(39Apa) und aus der *E. coli* MG1655-Tev2a Kultur mit dem Plasmid pSB16.23c102 $0,34 \pm 0,008$ mg/L•OD eGFP(39Apa) gereinigt. Damit war bei Verwendung der *E. coli* MG1655-Tev2a Zellen und der Ko-Expression der TEV-Protease eine Steigerung der Ausbeute um den Faktor 1,9 möglich. In diesem Rahmen belief sich auch die durch die Expression von eGFP(39THF-K) zusammen mit der TEV-Protease in *E. coli* MG1655-Tev2a gesteigerte Ausbeute. Aus der MG1655 Kultur, die das Plasmid pSB16.100d2 trug, wurde

0,23 ± 0,001 mg/L•OD eGFP(39THF-K) und aus der MG1655-Tev2a Kultur, die das Plasmid pSB16.100d102 trug, wurde 0,50 ± 0,01 mg/L•OD eGFP(39THF-K) isoliert. Mittels ESI-Massenspektrometrie wurde die genaue Zusammensetzung von eGFP(39Apa) (Abbildung 46 C und D) und eGFP(39THF-K) (Abbildung 46 E und F) bestätigt.

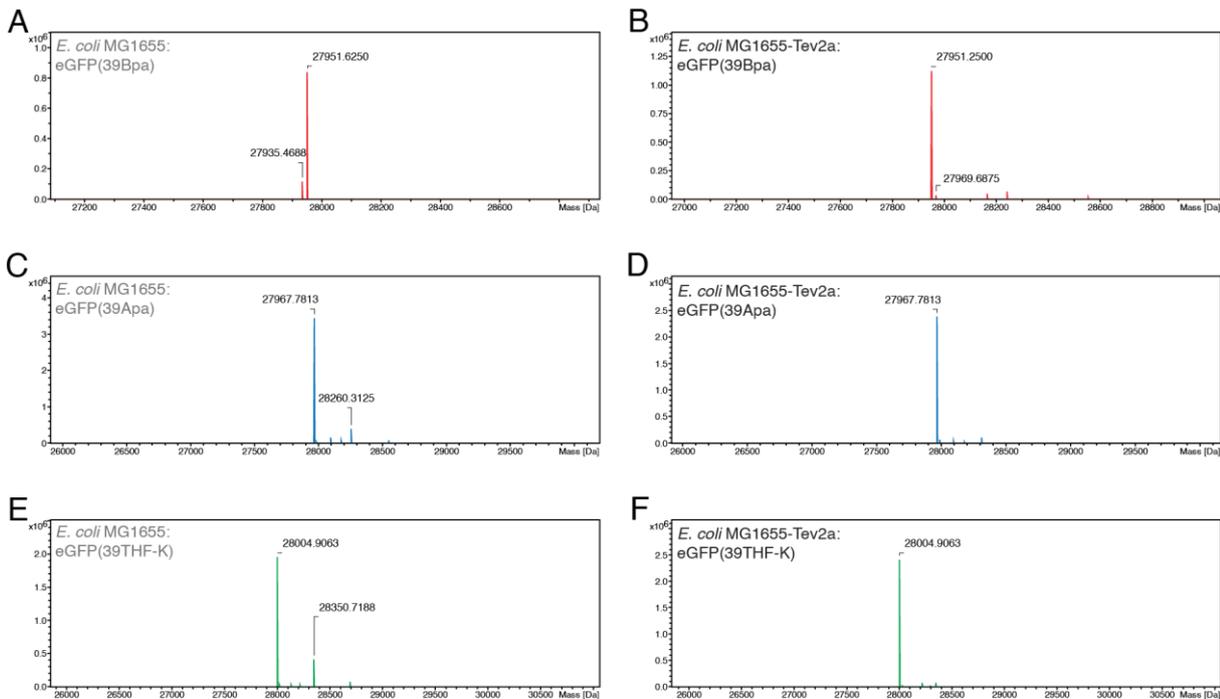


Abbildung 46: ESI-Massenspektrum des gereinigten eGFP(39a), exprimiert in *E. coli* MG1655 oder zusammen mit der TEV-Protease in *E. coli* MG1655-Tev2a. (A) eGFP(39Bpa) exprimiert in *E. coli* MG1655. $MW_{\text{kalk.}} = 27968,5$ Da; $MW_{\text{obs.1}} = 27951,6$ Da; $\Delta m_1 = -17,0$ Da, Abspaltung eines Wassermoleküls in Folge einer Esterbindung der Bpa-Seitenkette mit benachbarten Aminosäure-Seitenketten; $MW_{\text{obs.2}} = 27935,4$ Da; $\Delta m_2 = -35,1$ Da, Abspaltung von zwei Wassermolekülen in Folge einer Diesterbindung der Bpa-Seitenkette mit benachbarten Aminosäure-Seitenketten. (B) eGFP(39Bpa) exprimiert als Bicistron in *E. coli* MG1655-Tev2a. $MW_{\text{kalk.}} = 27968,5$ Da; $MW_{\text{obs.1}} = 27951,3$ Da; $\Delta m_1 = -17,6$ Da, Abspaltung eines Wassermoleküls in Folge einer Esterbindung der Bpa-Seitenkette mit benachbarten Aminosäure-Seitenketten; $MW_{\text{obs.2}} = 27969,7$ Da; $\Delta m_2 = -0,8$ Da. (C) eGFP(39Apa) exprimiert in *E. coli* MG1655. $MW_{\text{kalk.}} = 27966,7$ Da; $MW_{\text{obs.1}} = 27967,8$ Da, $\Delta m_1 = +1,1$ Da; $MW_{\text{obs.2}} = 28360,3$ Da; $\Delta m_2 = +393,6$ Da. (D) eGFP(39Apa) exprimiert mit der TEV-Protease in *E. coli* MG1655-Tev2a. $MW_{\text{kalk.}} = 27966,7$ Da; $MW_{\text{obs.1}} = 27967,8$ Da; $\Delta m_1 = +1,1$ Da. (E) eGFP(39THF-K) exprimiert in *E. coli* MG1655. $MW_{\text{kalk.}} = 28003,8$ Da; $MW_{\text{obs.1}} = 28004,9$ Da; $\Delta m_1 = +1,1$ Da; $MW_{\text{obs.2}} = 28350,7$ Da; $\Delta m_2 = +346,9$ Da. (F) eGFP(39THF-K) exprimiert mit der TEV-Protease in *E. coli* MG1655-Tev2a. $MW_{\text{kalk.}} = 28003,8$ Da; $MW_{\text{obs.1}} = 28004,9$ Da; $\Delta m_1 = +1,1$ Da.

3.3.2.5 Fusion von TEV-Protease mit dem Sumo-Protein zu Steigerung der Löslichkeit Aktivität

Die TEV-Protease ist eine aus dem *Tobacco Etch Virus* stammende sequenzspezifische Cysteinprotease (Kapust *et al.*, 2001). Da die rekombinante Expression der TEV-Protease mit geringerer Löslichkeit verbunden ist, erfolgt die Produktion der TEV-Protease meist als Fusionsprotein mit dem Maltose Binding Protein (MBP) (Kapust & Waugh, 1999). Um herauszufinden, ob eine löslichere und dadurch aktivere TEV-Protease die Suppression des Amber-Stop Codons in dem *E. coli* Stamm MG1655-Tev2a positiv beeinflusst, wurde die TEV-Protease als Fusion mit dem Sumo Protein exprimiert (Sadr *et al.*, 2017). Das Sumo Protein

wurde aufgrund der im Vergleich zum MBP geringeren Größe gewählt, um die Plasmidgröße nur wenig zu erhöhen. Die kodierende Sequenz des Fusionsproteins aus TEV-Protease und Sumo wurde mittels PCR generiert. Das Plasmid pSUMO#1 (erhalten von Michaela Gebauer, Lehrstuhl für Biologische Chemie) wurde hierbei als Matrizen-DNA für die PCR-Amplifikation mit den *Primern* rfCS48 und rfCS49 verwendet. Durch den *Primer* rfCS37 wurde die Erkennungssequenz für das Restriktionsenzym *HindIII*, die Sequenz für die OmpA-RBS wie auch das Start-ATG am 5'-Ende angefügt. Die Sequenz der TEV-Protease wurde durch PCR unter Verwendung des Plasmides pSB16.6c102 als Matrizen-DNA und den *Primern* rfCS50 und rfCS51 amplifiziert. Die *Primer* rfCS49 und rfCS50 wiesen eine 18 bp-lange, zueinander homologe Sequenz auf, wodurch die beiden zuvor generierten PCR-Produkte im zweiten Schritt mittels einer *Overlap-Extension* PCR zu einem DNA-Fragment assembliert werden konnten. Ein Verdau des daraus resultierenden PCR-Produktes mit *HindIII* ermöglichte die Klonierung in die ebenfalls mit *HindIII* verdauten und zusätzlich dephosphorylierten Vektoren pSB16.6c102, pSB16.23c102 und pSB16.100d102. Die daraus resultierenden Plasmide, die zusätzlich zu der für die Erkennung der nicht-natürlichen Aminosäuren benötigten Aminoacyl-tRNA Synthetasen BpaRS, ApaRS oder PylRS, die für die Fusion aus TEV-Protease und Sumo-Protein kodierten, wurden pSB16.6c110, pSB16.23c110 und pSB16.100d110 genannt.

Im Anschluss an die Sequenzierung der klonierten Plasmide wurde mit diesen der *E. coli* Stamm MG1655-Tev2a transformiert. Der Effekt der TEV-Protease als Fusionsprotein auf die Suppression des Amber-Stop Codons durch die Proteolyse des Rf-1 wurde wiederum über die Fluoreszenzintensität des Reporterproteins eGFP(39a) evaluiert. Dabei dienten Kulturen von *E. coli* MG1655-Tev2a, welche die Plasmide pSB16.6c2, pSB16.6c102, pSB16.23c2, pSB16.23c102, pSB16.100d2 oder pSB16.100d102 trugen, sowie *E. coli* MG1655, welche die Plasmide pSB16.6c2, pSB16.23c2 oder pSB16.100d2 trugen, als Vergleich. Bei einer $OD_{550} = 0,4$ wurden die entsprechenden nicht-natürliche Aminosäuren zugegeben und nach einer Stunde die Expression des Reporterproteins mit aTc induziert. Die Kulturen von *E. coli* MG1655-Tev2a, die eGFP(39a) zusammen mit dem Fusionsprotein aus TEV-Protease und Sumo exprimierten, zeigten kein stärkeres Fluoreszenzsignal als die *E. coli* MG1655-Tev2a Kulturen, bei denen die TEV-Protease ohne den Fusionspartner ko-exprimiert wurde (Abbildung 47). Zudem konnten keine Hinweise auf negative Auswirkungen hinsichtlich der Lebensfähigkeit der Zellen aufgrund einer möglichen Aktivitätssteigerung der TEV-Protease beobachtet werden. Weder eine verminderte Zelldichte noch die für tote Zellen typische Fluoreszenzsignale, lieferten einen Hinweis dahingehend. Zusammengefasst führt die Ko-Expression der TEV-Protease als Fusionsprotein mit SUMO nicht zu einer weiter verbesserten Suppressionseffizienz mit dem *E. coli* Stamm MG1655-Tev2a.

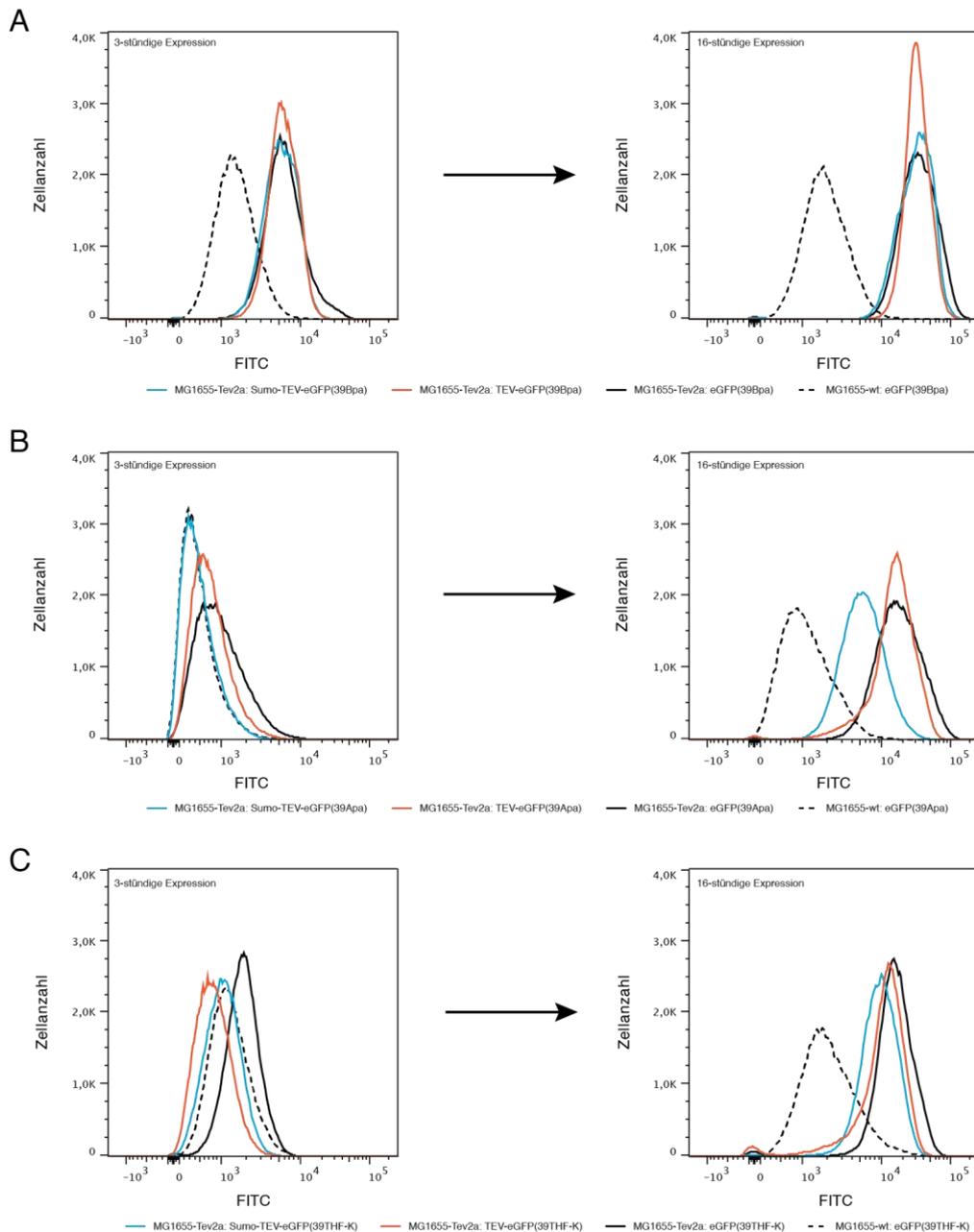


Abbildung 47: Auswirkung des Fusionsproteins aus TEV-Protease und Sumo auf die Suppression des Amber-Stop Codons in dem Stamm *E. coli* MG1655-Tev2a. Verglichen wurde das Fluoreszenzsignal, im Fall der Expression von eGFP(39a) in *E. coli* MG1655 (schwarz, unterbrochene Linie) oder MG1655-Tev2a (schwarz), die Ko-Expression von eGFP(39a) mit der TEV-Protease (orange) und die Ko-Expression von eGFP(39a) mit dem Fusionsprotein aus TEV-Protease und Sumo (blau) in *E. coli* MG1655-Tev2a. Dies wurde jeweils in Kombination mit den ebenfalls auf den Plasmiden kodierten Aminoacyl-tRNA Synthetasen, BpaRS (A), ApaRS (B) oder PylRS(C) untersucht.

3.3.2.6 Evaluierung der maximal erreichbaren Suppression des Amber-Stop Codons mittels der *E. coli* Suppressor-tRNA SupE

Wie in Abschnitt 3.3.2.3 bereits beschrieben, wurde vermutet, dass die Steigerung der Fluoreszenzintensität aufgrund der verbesserten eGFP(39a)-Expression von der für die Suppression des Amber-Stop Codons verwendeten durch Protein-Engineering generierten

Aminoacyl-tRNA Synthetase beziehungsweise von der eingebauten nicht-natürlichen Aminosäure abhängt. Um herauszufinden in welchem Ausmaß die Expression von eGFP(39a) durch die Verwendung von *E. coli* MG1655-Tev2a in Kombination mit der TEV-Protease im Vergleich zu der Expression in *E. coli* MG1655, unter idealen Bedingungen mit einer natürlichen Aminoacyl-tRNA Synthetase, gesteigert werden kann, wurde eine endogene Suppressor-tRNA genutzt. Verwendet wurde dafür die aus dem *E. coli* K12-Stamm XL-1 Blue (Bullock, 1987) stammende, endogene Suppressor-tRNA SupE, welche in der Literatur auch glnV44 oder su⁺² genannt wird (Inokuchi *et al.*, 1979) und für das Amber-Stop Codons spezifisch ist. SupE wird von der *E. coli* endogenen Glutaminyl-tRNA Synthetase mit Gln beladen und führt dadurch bei der Transkription zum Einbau von Gln in die Polypeptidkette anstelle der Termination am Amber-Stop Codon.

Für die Evaluierung der Expressionssteigerung bei der Suppression des Amber-Stop Codons in dem *E. coli* Stamm MG1655-Tev2a wurde die auf den Plasmiden pSB16.6c2, pSB16.6c102 und pSB16.6c110 kodierte künstliche Suppressor-tRNA (Wang & Schultz, 2001) gegen die Suppressor-tRNA SupE ausgetauscht. Die DNA-Sequenz von SupE wurde aus den vier *Oligonukleotidprimern* rfCS54, rfCS55, rfCS56 und rfCS57 assembliert, indem diese erst mit Hitze denaturiert wurden und durch langsames Abkühlen zu einem doppelsträngigen DNA-Fragment hybridisierten. Dieses DNA-Fragment konnte über die zum Schnittmuster der Restriktionsenzyme *Pst*I und *Eco*RI komplementären *Sticky-Ends* in die mit *Pst*I und *Eco*RI verdauten Vektoren inseriert werden. Zusätzlich wurde die ebenfalls auf den Plasmiden kodierte BpaRS, durch Deletion eines Großteils der kodierenden Region, inaktiviert. Dafür wurden die zuvor klonierten Plasmide mit dem Enzym *Eam*I verdaut, welches mit zwei Schnittstellen das aktive Zentrum flankiert, wobei ein Fragment von 606 bp ausgeschnitten wurde und mit einer 27-bp langen Sequenz, generiert aus zwei hybridisierten *Oligonukleotidprimern* LibCS9 und LibCS10, ligiert. Die resultierenden Plasmide kodieren alle für die Suppressor-tRNA SupE und für das Reporterprotein eGFP(39a) (pSB16.110g2), eGFP(39a) zusammen mit der TEV-Protease (pSB16.110g102) oder eGFP(39a) zusammen mit dem Fusionsprotein aus Sumo und TEV-Protease (pSB16.110g110). Im Anschluss an die Überprüfung mittels Sequenzierung wurden die drei Plasmide jeweils in den *E. coli* Stamm MG1655-Tev2a transfiziert und zusätzlich wurde *E. coli* MG1655 mit dem Plasmid pSB16.110g2 transformiert.

Um die maximale Expressionssteigerung bei der Suppression des Amber-Stop Codon in dem *E. coli* Stamm MG1655-Tev2a zu evaluieren, wurde wieder die Fluoreszenzintensität des Reporterproteins eGFP(39a) mit dem FACS-Gerät detektiert und verglichen. Die Kultivierung erfolgte ausgehend von jeweils einem Einzelklon in einer Flüssigkultur. Bei einer OD₅₅₀ = 0,6 wurde die Expression des Reporterproteins mit aTc für 3 Stunden beziehungsweise 16 Stunden induziert.

Nach 3-stündiger Expression des Reporterproteins konnte bereits ein Unterschied zwischen den beiden *E. coli* Stämmen MG1655-*Tev2a* und MG1655 festgestellt werden (Abbildung 48). Bei Verwendung desselben Plasmids für eGFP(39Gln), pSB16.110g2, wurde bei der Expression in *E. coli* MG1655-*Tev2a* eine Steigerung der Fluoreszenzintensität gegenüber MG1655 um den Faktor 5,2 detektiert (Abbildung 48, schwarze Linien). Die zusätzliche Ko-Expression der TEV-Protease hatte zu diesem Zeitpunkt noch keine signifikanten Auswirkungen auf die Expression des Reporterproteins, was sich jedoch nach 16-stündiger Expression änderte und in einer Verschiebung des Fluoreszenzsignals nach rechts (Abbildung 48, rote und blaue Linien) äußerte. In diesem Stamm wurde unter Verwendung der beiden für die TEV-Protease kodierenden Plasmide, pSB16.110g102 oder pSB16.110g110, eine um den Faktor 7,5 beziehungsweise 11,3 gesteigerte Fluoreszenzintensität im Vergleich zu der Expression mit pSB16.110g2 detektiert (Abbildung 48, rechts). Dabei fiel auf, dass sich im Gegensatz zu den zuvor durchgeführten Experimenten aus Abschnitt 3.3.2.5 die Ko-Expression der Fusion aus TEV-Protein und Sumo-Protein stärker positiv auf das eGFP(39a)-Signal auswirkte als die Ko-Expression der TEV-Protease ohne Fusionspartner.

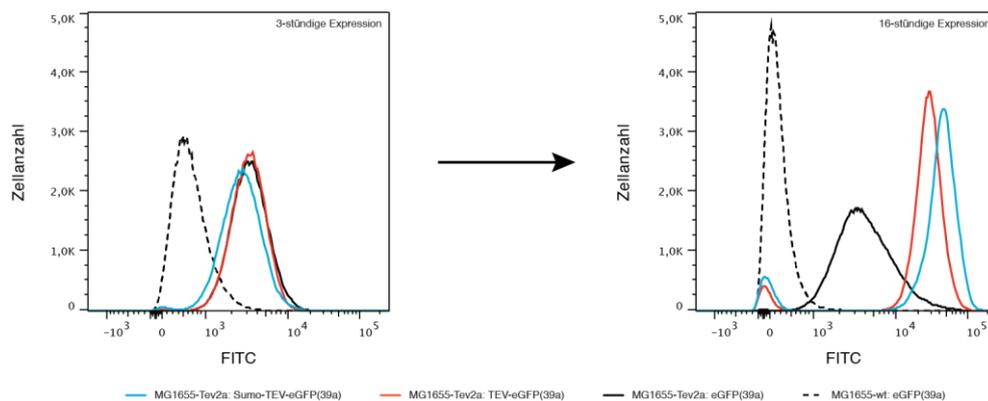


Abbildung 48: Suppression des Amber-Stop Codons mit der Suppressor-tRNA *SupE* bei fakultativ inaktiviertem *Rf-1*. Dargestellt sind die Fluoreszenzsignale, welche nach Expression von eGFP(39a) in *E. coli* MG1655 (schwarz, unterbrochene Linie) oder MG1655-*Tev2a* (schwarz), die Ko-Expression von eGFP(39a) und TEV-Protease (orange) und die Expression von eGFP(39a) zusammen mit dem Fusionsprotein aus TEV-Protease und Sumo (blau) in *E. coli* MG1655-*Tev2a* mit dem FACS-Gerät detektiert wurden.

Unerwarteterweise war allerdings nach 16-stündiger eGFP(39a)-Expression in *E. coli* MG1655 das Fluoreszenzsignal um den Faktor 2,8 geringer als das Signal derselben Kultur nach 3-stündiger Expression (Abbildung 48, schwarze unterbrochene Linien). Daraus resultierte bei einem Vergleich der beiden Expressionen — *E. coli* MG1655-*Tev2a* mit pSB16.110g102 oder pSB16.110g110 beziehungsweise *E. coli* MG1655 mit pSB16.110g2 — nach 16 Stunden ein um den Faktor 130 bis 198 gesteigertes Fluoreszenzsignal (Abbildung 48, rechts). Eine Erklärung, warum der *E. coli* Stamm MG1655-*Tev2a* deutlich besser mit der Suppression des Amber-Stop Codon durch *SupE* zurechtkam als der *E. coli* Stamm MG1655, gibt es nicht. Grundsätzlich lässt dieses Experiment zwei Schlussfolgerungen zu. Erstens scheint der *Rf-1*

durch die Einführung der TEV-Proteaseschnittstelle an Aktivität verloren zu haben, da in dessen Anwesenheit die Suppressionseffizienz von SupE erhöht ist. Zweitens führt die Ko-Expression der TEV-Protease und mehr noch ihres Fusionsproteins mit SUMO zu einer deutlichen Verstärkung dieses Effektes, was das Konzept der fakultativen Inaktivierung des Rf-1 zur Verbesserung der Suppression im Prinzip bestätigt.

4 Diskussion

4.1 Synthetische Biologie: Die Zukunft des Protein-Engineerings

In den 80er Jahren begann mit der Produktion von rekombinanten Proteinen eine Erfolgsgeschichte, die noch bis heute fortgesetzt wird. Sie beruht auf dem Einbringen von Fremd-DNA in ein Bakterium sowie der gezielten Manipulation der regulatorischen Elemente, um effiziente Transkription und Translation des Proteins durch den Wirtsorganismus zu bewirken. In den Anfängen fokussierten sich Forscher auf die Produktion von naturidentischen Proteinen für die Rote wie auch für die Weiße Biotechnologie (Riggs *et al.*, 1980), das in eine weitere Technologie mündete, die als Protein-Engineering bekannt ist. Diese Technologie beschreibt die serielle Veränderung der Eigenschaften durch Eingriffe in die der Struktur des Proteins (Ulmer, 1983). Die Grundlage dafür bildete einerseits die steigende Anzahl an gelösten Kristallstrukturen und andererseits die einfachere Manipulierbarkeit der für das rekombinante Protein kodierenden DNA-Sequenz. So wurde es durch *in vitro* Mutagenese möglich, gezielt Insertionen, Deletionen und Substitutionen vorzunehmen (Ulmer, 1983).

Mittlerweile ist es üblich, ganze Varianten-Bibliotheken der für das Protein kodierenden Sequenz mittels Sättigungsmutagenese, zufälliger Mutagenese oder Gen-*Shuffling* zu generieren und hinsichtlich der gewünschten Eigenschaften zu selektieren (Dalby, 2003). So konnten zum Beispiel bereits Enzyme mit gesteigerter Stabilität, unter anderem für den Einsatz in der Biokatalyse, therapeutische Proteine für Substitutionstherapien, Antikörper oder -fragmente wie auch alternative Bindeproteine generiert werden (Parren & Lugovskoy, 2013). Der nächste Schritt beim Protein-Engineering ist die Konstruktion von Proteinen oder Biopolymeren mit neuartigen Eigenschaften, die mit dem Satz an 20 natürlichen Aminosäuren nur schwer oder gar nicht realisiert werden können. Der Einbau von nicht-natürlichen Aminosäuren bietet dabei stetig zunehmende Möglichkeiten zur Funktionalisierung von Proteinen (Jin *et al.*, 2019). So ist mittlerweile die Erweiterung des genetischen Codes in der pharmazeutischen Industrie angekommen, wobei sich neuartige Therapeutika noch in der prä-klinischen oder auch in der klinischen Entwicklung befinden (Kang *et al.*, 2018). Die Firma Intrexon hat beispielsweise in einer Kollaboration mit der Firma Orgnics (Florida, USA) durch Einbau von nicht-natürlichen Aminosäuren ein neuartiges Lantibiotikum (OG716) entwickelt, um eine Infektion mit dem pathogenen *C. difficile* zu behandeln (Kers *et al.*, 2018). Dasselbe Ziel verfolgt die Firma Novacta Biosystems (Welwyn Garden City, UK) mit ihrem Lantibiotik-Derivat NVB302, das bereits die klinische Phase 1 bestanden hat (Ongey *et al.*, 2017).

Die Firma Ambrx verfolgt mit ihrer auf Säugetierzellen basierenden Produktionsplattform EuCODE™ die positionspezifische Funktionalisierung von therapeutischen Proteinen (Tian *et al.*, 2014). Sie haben den Her2-Antikörper, der den auf Karzinomen Zellen vorkommenden

Her2-Rezeptor bindet, über den Einbau einer nicht-natürlichen Aminosäure mit dem cytotoxischen Tubulin Inhibitor Amberstatin269 gekoppelt. Dieser α HER1-ADC (Antibody Drug Conjugate) ARX788 wird gerade in der klinischen Phase 1 evaluiert.

4.2 Borocaline als Biologika für die Diagnostik und Therapie von Krebserkrankungen und Virusinfektionen

Die Funktionalisierung von Proteinen mit nicht-natürlichen Aminosäuren bietet weitaus mehr Möglichkeiten als die gezielte Konjugation von Antikörpern. So ist es gelungen, in ein Lipocalin (Lcn2) eine Boronophenyl-Seitenkette einzuführen, um ein synthetisches zuckerbindendes Protein zu generieren (Edwardraja *et al.*, 2017). Kohlenhydratstrukturen sind attraktive Ziele für die Entwicklung von neutralisierenden Antikörper — oder alternativen Bindeproteinen — gegenüber Viren, die auf der Oberfläche des Virons präsentiert werden und eine entscheidende Rolle beim Eintritt in die Wirtszelle spielen (Murin *et al.*, 2019). Es wurden bereits erfolgreich Antikörper gegen das HI-Virus (Falkowska *et al.*, 2014) und das Dengue-Virus (Rouvinski *et al.*, 2015) entwickelt. Aufgrund der aktuellen Covid19 Pandemie stehen auch die auf dem Spikeprotein vorkommenden Glycanstrukturen des Virus SARS-Cov19 im Fokus der Wissenschaftler (Pinto *et al.*, 2020).

Obwohl Viren die Glycosylierungsmaschinerie der Wirtszelle benutzen, um durch Maskierung von Proteinstrukturen dem Immunsystem des Wirtsorganismus zu entkommen, unterscheiden sich die Glycanstrukturen auf den Vironen von den endogenen Oligosacchariden (Wang *et al.*, 2015). Auch bei Krebszellen unterscheidet sich das Glycosylierungsmuster von gesunden Zellen. So wird das Glycan LewisY auf Karzinomen durch die Expression auf Glycoproteinen und Glycosphingolipiden auf der Zelloberfläche präsentiert. LewisY spielt eine Rolle für das Tumorzellwachstum, die Adhäsion, Metastasierung und Proliferation (Hakomori, 1985). Der Antikörper BR96 gegen diese Glycanstruktur erreichte zwar klinische Studien, konnte dabei aber keine ausreichende Immunantwort mit wirksamer Toxizität für die Krebszellen vermitteln (Saleh *et al.*, 2000; Scott *et al.*, 2007). Das vorrangige Problem von Antikörper gegen Zuckerstrukturen ist die Kreuzreaktivität mit nicht pathologischen Glycanstrukturen. Im Fall von LewisY konnten die bisher entwickelten Antikörper auch andere Lewis-Zuckerstrukturen binden (Noble *et al.*, 2013). Eine solche Kreuzreaktivität kann — im schlimmsten Fall — eine Autoimmunerkrankung auslösen (Kappler & Hennet, 2020b). Deshalb ist eine Entwicklung von alternativen Bindeproteinen gegen krankheitsrelevante Zuckerstrukturen von enormer Bedeutung.

Die in dieser Arbeit untersuchten alternativen Bindeproteine, die Borocaline, ermöglichen durch die Funktionalisierung mit einer Boronsäure die kovalente Bindung von *cis*-Diolen, wie sie häufig bei Zuckern vorkommen und bieten daher einen attraktiven Ansatz für die Weiterentwicklung. Die Boronsäure unterscheidet zwar kaum zwischen verschiedenen *cis*-Diol

tragenden Zuckermolekülen, dafür eignet sich aber das *Scaffold*-Protein, das humane Lipocalin Lcn2, hervorragend für die Generierung einer substratspezifischen Bindetasche durch konventionelles Protein-Engineering. Die strukturelle Grundlage dafür bieten die vier Schleifen der Bindetasche, in denen 20 Aminosäurereste mittels ortsgerichteter Sättigungsmutagenese randomisiert und anschließend auf eine passende Bindetasche selektiert werden kann (Richter *et al.*, 2014). Diese Herangehensweise erlaubte es in der Vergangenheit bereits, eine Vielzahl von Lcn2-*Scaffolds* hinsichtlich der Bindung vorgegebener Moleküle zu selektieren, die unter dem Markennamen Anticaline bekannt sind. Solche Anticaline gibt es bereits für proteinogene Strukturen wie ED-B, Hsp70, VEGFR-3 oder PSMA, Peptide (Eichinger *et al.*, 2022) aber auch für biomedizinisch relevante Haptene wie Colchicin, oder Petrobactin (Rothe & Skerra, 2018; Deuschle *et al.*, 2020). Zu den ausgeprägten Bindungseigenschaften dieser Anticaline kommt eine niedrige Immunogenität, steuerbare Retentionszeit im Blutplasma, effiziente Produktion in *E. coli* wie auch eine gute Gewebepenetration sowie die Möglichkeit, die Anticaline durch Fusion oder chemisch Konjugation mit zusätzlichen Funktionen auszustatten (Gebauer & Skerra, 2012), was sowohl bei Krebserkrankungen als auch Virusinfektionen von Belang ist.

Die Borocaline sind zwar noch nicht so weit entwickelt, um komplexe Glycanstrukturen zu erkennen, jedoch wurde die Bindung einfacher Zuckermoleküle und Zucker-Derivate mit moderaten Affinitäten (K_D -Bereich: mM bis μ M) demonstriert. Ein Engineering der Bindetasche des Lcn2 resultierte nicht in einer Verbesserung der Affinität der Borocalins gegenüber den komplexen Zuckermolekülen. Vermutet wurde ein Problem beim verwendeten Boronsäure-Anker, dem Bpa, der mittels Amber-Stop Codon-Technologie in die Polypeptidkette eingebaut wurde.

Boronsäuren können in zwei Konfigurationen vorliegen, dem trigonal-planaren Zustand oder als Tetraeder. Die tetraedrische Form wird mit dem freien Valenzorbital unter Reaktion mit einem OH^- bei höherem pH-Wert gebildet (Pizer & Babcock, 1977). Die nicht-natürliche Aminosäure Bpa hat einen $\text{p}K_S$ -Wert von ~ 9 , das bedeutet, dass bei physiologischem pH-Wert der Großteil in der trigonal-planaren Struktur vorliegt. Das wurde auch mit der Kristallstruktur des Apo-Borocalins, bei dem der Kristall bei einem pH-Wert von 7,5 gezüchtet wurde, bestätigt (Edwardraja *et al.*, 2017). Die tetraedrische Struktur der Bpa-Seitenkette konnte dagegen in der Kristallstruktur des Borocalin/Nitrocatechol-Komplexes nachgewiesen werden. In diesem Fall wurde der Kristall für bei pH 8,5 gezüchtet, wobei zirka 20 % des Boronsäurerestes als Boronat in tetraedrischer Struktur vorliegen sollte. In der Kristallstruktur kann man zudem eine polare Interaktion einer Asn-Seitenkette erkennen, die, vermutlich nach Ausbildung des Tetraeders, diesen stabilisiert. 4-Nitrocatechol ist ein besonders guter Bindungspartner für den Boronsäurerest aufgrund der elektronenziehenden Eigenschaften der Nitrogruppe und der nahezu planaren Struktur des gebildeten Fünfrings (Sommer *et al.*, 2020).

Erst die Kristallstruktur konnte belegen, dass das Boronat gegenüber der Boronsäure als Reaktionspartner mit dem *cis*-Diol bevorzugt wird und daher eine reaktivere Boronsäure, als nicht-natürliche Aminosäure erfolgsversprechender hinsichtlich der kovalenten Komplexbildung mit Zuckermolekülen ist.

Das von Dr. Markus Anneser synthetisierte fluorierte Bpa-Derivat, Bfa, weist genau diese Eigenschaften auf. Die Einführung von zwei Fluor-Atomen in *ortho*-Position des Phenylrings von Bpa verringert den pK_s . Für den Einbau dieser nicht-natürlichen Aminosäure, Bfa, mit der für Bpa spezifischen Aminoacyl-tRNA Synthetase BpaRS sollten diese Substitutionen aufgrund des geringen Unterschiedes des Atomradius zum Wasserstoff keine entscheidende Rolle spielen. Jedoch wurde das Säure/Basen-Gleichgewicht verschoben und die Boronsäure-Seitenkette liegt nun vorwiegend als anionischer Tetraeder vor. Die molekulare Erkennung von Bfa durch BpaRS wird also nicht nur durch die strukturellen Unterschiede der Boronsäure verhindert, sondern durch die negative Ladung. Zwar war es möglich eGFP(39Bfa) mit geringer Ausbeute zu exprimieren, (Abbildung 21) was daran liegt, dass trotz der Verschiebung des Gleichgewichtes noch ein geringer Anteil der Bfa Seitenkette trigonal-planar vorliegt. Dementsprechend muss für jede nicht-natürliche Aminosäure eine spezifische Aminoacyl-tRNA Synthetase generiert werden, was einen großen Zeitaufwand mit sich bringt. Eine weitere Frage, die eine reaktivere borhaltige nicht-natürliche Aminosäure aufwirft, ist die nach der Interaktion mit den in der *E. coli* Zelle vorkommenden *cis*-Diolen.

4.3 Herausforderung bei der Selektion spezifischer Aminoacyl-tRNA Synthetasen für nicht-natürliche Aminosäuren

Die größte Herausforderung bei der Selektion von neuen Aminoacyl-tRNA Synthetasen für nicht-natürliche Aminosäuren ist die Planung und Generierung einer geeigneten Gen-Bibliothek. Dafür muss im ersten Schritt ein orthogonales Paar aus Aminoacyl-tRNA Synthetase und Suppressor-tRNA ausgewählt werden. Zumeist wurden Aminoacyl-tRNA Synthetasen für sterisch anspruchsvolle Seitenketten ausgehend von der TyrRS aus *Methanocaldococcus jannaschii* und für langgestreckte Seitenketten ausgehend von der PylRS aus *Methanosarcina barkeri* oder *Methanosarcina mazei* selektiert. In diese Arbeit wurde versucht eine neue Aminoacyl-tRNA Synthetase ausgehend von der *Mj*TyrRS zu generieren. Bei der Analyse von bereits publizierten Aminoacyl-tRNA Synthetasen für nicht-natürliche Aminosäuren mit bekannter Raumstruktur, die ebenfalls ausgehend von der *Mj*TyrRS engineeriert wurden, fiel auf, dass die sechs Aminosäurepositionen Tyr³², Leu⁶⁵, Gln¹⁰⁹, Asp¹⁵⁸, Ile¹⁵⁹, Leu¹⁶² zentral für die meisten Gen-Bibliotheken zur Selektion neuer Aminoacyl-tRNA Synthetase-Spezifitäten waren (Krahn *et al.*, 2020). Das liegt zum einen daran, dass Tyr³² und Asp¹⁵⁸ auch in die selektive Erkennung der Aminosäure-Seitenkette des nativen Substrates, Tyr, involviert sind (Kobayashi *et al.*, 2003). Hinzu kommt, dass die erste engineerierte Aminoacyl-

tRNA Synthetase, die die nicht-natürliche Aminosäure *O*-Methyl-L-tyrosin erkennt und ausgehend von der *Mj*TyrRS generiert wurde, Mutationen an den Aminosäurepositionen 32, 158, 107 und 162 aufwies (Wang *et al.*, 2001).

Bei der Konstruktion einer geeigneten Gen-Bibliothek ist darauf zu achten, dass die Anzahl an randomisierten Aminosäurepositionen zu einer Komplexität führt, die transformierbar und selektierbar ist. Die maximale Transformationseffizienz von *E. coli* Zellen liegt in der Regel bei 10^9 - 10^{10} Transformanden pro μg DNA, was damit auch die maximale Komplexität der Gen-Bibliothek beschränkt. Bei der Selektion mit dem FACS-Instrument kann man eine solch komplexe Bibliothek jedoch nicht vollständig selektieren, da die Dauer der Selektion mit diesem Gerät die Anzahl an Varianten auf zirka 10^8 begrenzt. Allerdings kann man eine Vorselektion auf Agarplatten durchführen, um ausschließlich funktionsfähige Aminoacyl-tRNA Synthetasen zu erhalten. Bei dieser Vorselektion wird auf Überleben selektiert, da das Chloramphenicolresistenz vermittelnde Gen ein Amber-Stop Codon trägt und die Zellen auf Agarplatten mit diesem Antibiotikum ausplattiert werden. Umso entscheidender ist die Wahl des geeigneten Selektionsdrucks. Ein zu starker Selektionsdruck würde auch für schlechte Aminoacyl-tRNA Synthetasen letal sein, ein zu schwacher Selektionsdruck würde die Anzahl an mutierten Aminoacyl-tRNA Synthetasen nicht ausreichend stringent reduzieren, um die daraus folgende Selektion mit dem FACS-Gerät zu erleichtern. Allerdings werden bei einer solchen Vorselektion nicht nur Aminoacyl-tRNA Synthetasen angereichert, die die nicht-natürliche Aminosäure erkennen, sondern angesichts der nativen Funktion der *Mj*TyrRS vor allem auch jene, die Tyr erkennen. Der optimale Selektionsdruck wurde hier in einem Experiment mit einer bereits optimierten Aminoacyl-tRNA Synthetase (BpaRS) ermittelt. Dabei wurde unter anderem eine lineare Abnahme der Transformanden, im Einklang mit einer direktem Korrelation zu der verwendeten Chloramphenicol-Konzentration beobachtet.

Obwohl die Selektion aus der SlonoMax Gen-Bibliothek auf pSB16.6c2 mit und ohne Vorselektion durchgeführt wurde, konnten keine Aminoacyl-tRNA Synthetasen identifiziert werden, die Hca erkennen, sondern nur solche, die Tyr als Substrat verwenden. Das lag aber vor allem an der Verteilung der Aminosäurepositionen, die nach heutigem Kenntnisstand nicht ideal war. Ein strukturelles Alignment aller ausgehend von der *Mj*TyrRS engineerter Aminoacyl-tRNA Synthetase für die Erkennung von nicht-natürlichen Aminosäuren, zeigte dagegen, dass sterisch anspruchsvolle nicht-natürliche Aminosäuren wie Hca anders als erwartet im aktiven Zentrum gebunden werden. Deshalb wurde die in der DNA kodierte Information dieser bereits engineerter Aminoacyl-tRNA Synthetasen genutzt, um daraus eine fokussierte Gen-Bibliothek zu generieren. Hinzu kommt, dass die drei für die Konstruktion dieser Bibliothek mittels *Sexual*-PCR und Zufallsmutagenese gewählten Aminoacyl-tRNA Synthetasen Tyr nur noch schlecht als Substrat erkennen. Da diese Gen-Bibliothek jedoch nur eine begrenzte Diversität aufwies, wurde sie nach jeder Selektionsrunde durch Zufallsmutagenese erneut diversifiziert. Nach fünf Selektions- und Diversifizierungsrunden

konnte in der Einzelklonanalyse keine Aminoacyl-tRNA Synthetase selektiert werden, die Hca besser erkennt als Tyr. Allerdings ist nicht ausgeschlossen, dass einige dieser selektierten Mutanten die nicht-natürliche Aminosäure als Substrat akzeptiert. Eine fortgesetzte gerichtete Evolution könnte hierüber Aufschluss geben.

Ein er komplett anderer Ansatz wäre das Engineering einer Aminoacyl-tRNA Synthetase basierend auf der *MmPylRS*. Obwohl man bevorzugt *MjTyrRS* verwendet, um Aminoacyl-tRNA Synthetase für sterisch anspruchsvolle, nicht-natürliche Aminosäuren zu selektieren, wurde vor kurzem gezeigt, dass sich auch die *MmPylRS* dafür eignen würde (Tseng *et al.*, 2020). Der Vorteil hierbei wäre, dass das native Substrat der *MmPylRS*, Pyrrolysin, nicht zu den 20 natürlichen Aminosäuren gehört und daher weder im Medium enthalten ist, noch von *E. coli* synthetisiert wird.

Vielversprechender erscheint die Generierung einer Aminoacyl-tRNA Synthetase, die Bfa erkennt. Der strukturelle Unterschied von Bpa zu Bfa bei physiologischem pH-Wert bezieht sich vor allem auf die Boronatgruppe. Bei Bpa liegt diese trigonal-planar und bei Bfa als anionischer Tetraeder vor. Angesichts der ansonsten gegebenen strukturellen Ähnlichkeiten der Seitenkette würde sich für die Generierung einer Gen-Bibliothek zur Selektion einer geeigneten Aminoacyl-tRNA Synthetase für Bfa die Gen-Sequenz der BpaRS als Grundlage eignen.

Daran müsste man die bei der BpaRS direkt in die Substraterkennung involvierten Aminosäure-Positionen randomisieren. Vor allem das mit der trigonal-planaren Boronsäure von Bpa durch zwei Wasserstoffbrückenbindungen interagierende Glu¹⁶² würde im Fall der tetraedrisch vorliegenden anionischen Boronsäure zur elektrostatischen Abstoßung führen. Des Weiteren stellen die Aminosäuren Ser³² und Gln¹⁰⁹ ein Ziel für die Mutagenese dar, da diese über ein Wassermolekül mit der trigonal-planaren Boronsäure der Bpa interagieren und bei der tetraedrischen Konfiguration von Bfa vermutlich kein Platz für diese Wassermoleküle mehr wäre.

4.4 Limitierende Faktoren für die Proteinausbeute auf dem Weg zur Anwendbarkeit

Nach über 20 Jahren Amber-Stop Codon-Technologie wurden hoffnungsvolle Forschungsergebnisse hinsichtlich pharmazeutisch relevanter Biologika und proteinbiochemischer Neuerungen publiziert, eine industrielle Anwendung konnte dagegen nur im begrenzten Rahmen realisiert werden. Hierbei stehen vor allem geringe Proteinausbeuten bei der gentechnischen Herstellung den hohen Kosten für die benötigten nicht-natürlichen Aminosäuren gegenüber. Die Zugabe der nicht-natürlichen Aminosäuren in das Medium ist aktuell noch obligat, da es nur in Einzelfällen gelungen ist, Mikroorganismen für die *in vivo* Biosynthese dieser aus herkömmlichen Nährmediumsbestandteilen zu evolvieren (Zhang *et al.*, 2017).

Das Problem der geringen Proteinausbeute bei der Anwendung der Amber-Stop Codon-Technologie bleibt aber bestehen und lässt sich mit einer ungefähren Reduktion um den Faktor 10 im Vergleich zur rekombinanten Gen-Expression mit unterbrochenem Leserahmen beziffern. Neben der verminderten katalytischen Effizienz der engineereten Aminoacyl-tRNA Synthetasen (Beyer *et al.*, 2020) liegt es vor allem an der Konkurrenz des Rf-1 mit der Suppressor-tRNA um das Amber-Stop Codon. Um diese Konkurrenz zu unterbinden, wurde in dem *E. coli* B-Stamm NEB Express das für den Rf-1 kodierende *prfA*-Gen deletiert. Der erhaltene *E. coli* Stamm NEB Express Δ Rf-1 wurde hinsichtlich seiner Expressionseigenschaften im Vergleich zum ursprünglichen *E. coli* NEB Express evaluiert. Dafür wurde unter Verwendung der BpaRS und der nicht-natürlichen Aminosäure Bpa das Reporterprotein eGFP(39a) exprimiert und das durch Suppression des Amber-Stop Codons erzeugte, intakte fluoreszierende Protein detektiert. Die Expression in dem *E. coli* NEB Express Δ Rf-1 Stamm wies dabei eine um den Faktor 2,2 gesteigerte Fluoreszenz — gemessen mit einem Durchflusszytometer — im Vergleich zu *E. coli* NEB Express auf. Eine vergleichbare Steigerung konnte bei der Produktion und Reinigung des Reporterproteins ermittelt werden.

Eine andere Forschungsgruppe konnte durch Deletion des Rf-1 in *E. coli* BL21 (DE3) keine signifikante Steigerung der Fluoreszenz bei der Suppression eines Amber-Stop Codons in der kodierenden Sequenz des gleichen Reporterproteins detektieren (Johnson *et al.*, 2012). Jedoch konnten sie die generell hohe Proteinausbeute des *E. coli* BL21 (DE3)-Stammes durch diese Deletion bei der Suppression von zwei und drei Amber-Stop Codons steigern (Johnson *et al.*, 2012). Die Firma enGenes Biotech verfolgte einen anderen Ansatz, bei dem die Proteinproduktion von der Wachstumsphase entkoppelt ist. Dafür wurde der virale Polymerase-Inhibitor Gp2 in das Genom von *E. coli* BL21 (DE3) integriert, so dass mittels Induktion dieses Proteins die endogene Proteinproduktion und das Wachstum inhibiert werden kann. Die Expression der viralen T7 Polymerase wie auch die des Zielproteins, dessen Gen unter Kontrolle des viralen T7-Promotors steht, sind davon nicht betroffen. Mit dieser Expressionstrategie wurde im Vergleich zu der Expression in *E. coli* BL21 (DE3) die Proteinausbeute um den Faktor 2,3 gesteigert (Lemmerer *et al.*, 2019; Galindo Casas *et al.*, 2020). Der hier generierte *E. coli* NEB Express Δ Rf-1 Stamm ermöglicht dem gegenüber nicht nur gesteigerte Proteinausbeuten sondern auch eine einfache Handhabung ohne zusätzlich Induktion.

Eine Entkopplung von Proteinproduktions- und Wachstumsphase ist vor allem bei *E. coli* K12-Derivaten nötig, da diese im Gegensatz zu *E. coli* B-Stämmen eine Mutation im Rf-2 tragen, welche die Terminationseffizienz des Rf-2 um den Faktor 5 reduziert und dadurch eine Deletion des für den Rf-1 kodierenden *prfA*-Gens letal für die Zellen wäre (Gerdes *et al.*, 2003). In dieser Arbeit wurde der Rf-1 des *E. coli* K12 Stamm MG1655 mittels λ -Red Rekombination und Cas9 Selektion mutiert und so ein TEV-Protease-Erkennungsmotif in einer katalytisch

relevanten Domäne eingeführt. Damit sollte es möglich sein, den Rf-1 im Anschluss an die Wachstumsphase bei Induktion der Proteinproduktionsphase durch simultane Expression des Zielproteins und der TEV-Protease zu inaktivieren und so die Konkurrenz zwischen Suppressor-tRNA und Rf-1 um das Amber-Stop Codon zu beenden. Das dabei erhaltene Derivat *E. coli* MG1655-Tev2a wurde durch den Vergleich mit dem *E. coli* Stamm MG1655 in Kombination mit den Aminoacyl-tRNA Synthetasen BpaRS, ApaRS oder PylRS und in Gegenwart der korrespondierenden nicht-natürlichen Aminosäuren hinsichtlich der Expression eines Reporterproteins evaluiert. Das Reporterprotein eGFP(39a) trägt ein Amber-Stop Codon an einer unkritischen Stelle in der kodierenden Sequenz (Kuhn *et al.*, 2010), wobei dessen erfolgreiche Suppression zu einem fluoreszierenden Protein führt, welches mittels Durchflusszytometrie detektiert werden kann.

Die Ergebnisse zeigten, dass bei Verwendung des *E. coli* Stammes MG1655-Tev2a durch simultane Expression von eGFP(39a) und der TEV-Protease die Fluoreszenz im Vergleich zu *E. coli* MG1655 mit allen drei Aminoacyl-tRNA Synthetasen gesteigert werden konnte. Abhängig von der Aminoacyl-tRNA Synthetase, konnten aber Unterschiede festgestellt werden. Die Ko-Expression des eGFP(39THF-K) mit der TEV-Protease in *E. coli* MG1655-Tev2a und mit der PylRS zeigte aber auch, dass sich die Suppression des Amber-Stop Codons auf einen Teil der Zellpopulation negativ auswirkte.

Zusätzlich wurde das Reporterprotein im Schüttelkolben-Maßstab produziert und gereinigt. Auch hier wurde bei Verwendung des *E. coli* Stammes MG1655-Tev2a und simultaner Expression des Zielproteins mit der TEV-Protease die Proteinausbeute gesteigert. Allerdings korrelierte diese Steigerung nicht mit der mittels Durchflusszytometer detektierten Zunahme der Fluoreszenz.

Die durchgeführten Experimente deuteten auch darauf hin, dass die Mutationen, die für die Generierung des *E. coli* Stammes MG1655-Tev2a eingeführt wurden, selbst eine Auswirkung auf die Proteinexpression haben. Eine zusätzliche Inaktivierung des Rf-1 durch die TEV-Protease ist daher nicht immer notwendig. Der günstige Effekt eines inaktiviertem Rf-1 kommt umso weniger zu tragen je weniger effizient die verwendete Aminoacyl-tRNA Synthetase ist. So war es möglich, die Proteinausbeute ohne die Ko-Expression der TEV-Protease unter Verwendung des *E. coli* MG1655-Tev2a Stammes und in Kombination mit der ApaRS oder PylRS auf dasselbe Level zu steigern wie bei der Ko-Expression der TEV-Protease.

Für eine von der Aminoacyl-tRNA Synthetase unabhängige Evaluierung der fakultativen Inaktivierung des Rf-1 mit *E. coli* MG1655-Tev2a durch Ko-Expression der TEV-Protease, wurde die Suppressor-tRNA SupE von *E. coli* verwendet. Dieses Experiment zeigte, dass der *E. coli* Stamm MG1655-Tev2a die Suppression des Amber-Stop Codon durch SupE im Vergleich zu MG1655 förderte. Die eGFP(39a)-Expression in *E. coli* MG1655 wies eine

unerwartet niedrige Fluoreszenz auf. Deshalb ergab die Berechnung der Fluoreszenzsteigerung durch die Verwendung des *E. coli* Stammes MG1655-Tev2a den unerwartet hohen Faktor 130. Die fakultative Inaktivierung des RF-1 sollte trotz der noch bestehenden Unklarheiten Vorteile bringen in Richtung Anwendbarkeit des Einbaus nicht natürlicher Aminosäuren durch Suppression des Amber-Stop Codons.

5 Zusammenfassung

Ein zentrales Thema in der Synthetischen Biologie ist die Erweiterung des genetischen Codes. Dieser beruht auf 64 Triplett-Codons, die aus der Kombination der 4 DNA-Basen Adenin, Guanin, Cytosin und Thymin hervorgehen. Davon fungieren drei Codons als Stop-Codons, um die Translation zu terminieren, und die anderen 61 Codons sind in redundanter Weise den 20 Aminosäuren zugewiesen. Das bedeutet, dass einzelne proteinogene Aminosäuren von bis zu sechs verschiedenen Triplets kodiert werden. Die damit grundsätzlich gegebene Möglichkeit der Speicherung weiterer Informationen wird in der Natur teilweise durch die Erweiterung um eine zusätzliche Aminosäure, Selenocystein oder Pyrrolysin, realisiert. Für strukturelle Veränderungen am Protein nutzt die Natur demgegenüber fast ausschließlich die posttranslationale Modifikation, einen enzymatischen Prozess, der allerdings weniger strikt geregelt ist als die Proteinbiosynthese, bei der der genetische Code die Abfolge der Aminosäuren klar vorgibt.

Eine gezielte Modifizierung von Proteinen ist im Zeitalter der Gentechnologie durch die Erweiterung des genetischen Codes möglich, bei dem insbesondere verfügbare Stop-Codons nicht-natürlichen Aminosäuren neu zugeordnet werden. Bevorzugt verwendet man das seltenste Codon, das Amber-Stop Codon. Die ribosomale Translation dieses Codons erfolgt mit Hilfe eines für die nicht-natürliche Aminosäure spezifischen Paares aus Aminoacyl-tRNA Synthetase und Suppressor-tRNA, das die nicht-natürliche Aminosäuren erkennt, adenyliert und auf die für das Amber-Stop Codon spezifische Suppressor-tRNA lädt. Der Einbau in die Polypeptidkette erfolgt dann an dem Amber-Stop Codon, das zuvor in die kodierende Sequenz des Zielproteins an der gewünschten Stelle eingeführt wurde.

Eine solche Funktionalisierung wurde bei dem humanen Lipocalin 2 (Lcn2), einem natürlichen Bindeprotein, durchgeführt, wobei mittels Suppression des Amber-Stop Codons die nicht-natürliche Aminosäure Boronophenylalanin (Bpa) in die Polypeptidkette eingebaut wurde. Die so eingeführte Boronsäure kann mit *cis*-Diolen, wie sie vor allem in Zuckerstrukturen zu finden sind, zyklische Diester bilden. Dieses funktionalisierte Lcn2-Protein, Borocalin genannt, wie auch Protein-Varianten davon, zeigt Affinität gegenüber Zuckermolekülen mit K_D -Werten im mM bis μ M Bereich.

Die Kristallstruktur des Borocalins mit dem Modellmolekül 4-Nitrocatechol ermöglichte die Aufklärung des Komplexbildungsmechanismus dieser Boronsäure mit einem *cis*-Diol. Die Boronsäure, welche im trigonalen Zustand oder tetraedrisch vorliegen kann, bildet bei pH 8,5 den zyklischen Diester in anionischer Tetraederform. Da die Bpa-Seitenkette einen pK_S -Wert von ~ 9 hat, liegt bei physiologischem pH-Wert der Großteil jedoch in der weniger reaktiven trigonal-planaren Struktur vor. Für eine praktische Anwendung dieser Borocaline ist dagegen eine nicht-natürliche Aminosäure nötig, die auch unter physiologischen Bedingungen als

tetraedrisches Anion vorliegt, um die Affinität der funktionalisierten Bindeproteine gegenüber Zuckerstrukturen zu steigern.

Darüber hinaus wurde versucht, ausgehend von der *Mj*TyrRS eine neue Aminoacyl-tRNA Synthetase für den Einbau einer fluoreszierenden und sterisch anspruchsvollen nicht-natürlichen Aminosäure, 7-Hydroxycoumarylalanin (Hca), zu selektieren. Eine solche Selektion erwies sich aufgrund der nativen Funktion der *Mj*TyrRS, Tyr zu erkennen, als schwierig. Deshalb wurde ein neuartiger Ansatz für die Generierung einer solchen Aminoacyl-tRNA Synthetase etabliert. Dieser kombiniert nach dem Vorbild der Natur die Prozesse der Rekombination, Mutation und Selektion. Dafür wurden drei geeignete künstliche Aminoacyl-tRNA Synthetasen, die keine natürliche Aminosäure mehr erkennen und deren Substrat ähnliche sterische Eigenschaften hat, ausgewählt. Hierzu wurden die synthetischen homologen DNA-Sequenzen der Aminoacyl-tRNA Synthetasen HqAlaRS, AzAlaRS und Photo-DopaRS mittels *DNase*I verdaut und die daraus resultierenden Fragmente zu chimären Aminoacyl-tRNA Synthetasen durch PCR assembliert. Mit dieser Strategie sollte es künftig möglich sein, eine Aminoacyl-tRNA Synthetase für die sterisch anspruchsvolle nicht-natürliche Aminosäure Hca zu generieren.

Um das Problem der geringen Ausbeute von funktionalisiertem Protein bei der Suppression des Amber-Stop Codons zu lösen, wurden die Genome eines *E. coli* B-Stammes und eines *E. coli* K12-Stamm modifiziert. Da das Amber-Stop Codon normalerweise durch den Translationsterminationsfaktor Rf-1 erkannt wird, konkurriert letzterer mit der Suppressor-tRNA. Im Fall des *E. coli* B-Stammes war es möglich, das für den Rf-1 kodierende Gen zu deletieren. Beim *E. coli* K12-Stamm wurde der Rf-1 dahingehend mutiert, dass dieser eine TEV-Erkennungssequenz in einer katalytisch relevanten Domäne trägt und durch konditionale Ko-Expression der TEV-Protease posttranslational inaktiviert werden kann. Mit der permanenten wie auch der fakultativen Inaktivierung des Rf-1 konnte die Proteinausbeute abhängig von der eingesetzten Aminoacyl-tRNA Synthetase und der entsprechenden nicht-natürlichen Aminosäure um den Faktor 2,3 beziehungsweise 1,9 bis 4,4 gesteigert werden.

Die in dieser Arbeit gewonnenen Erkenntnisse bezüglich des Komplexbildungsmechanismus der Phenylboronsäuren, einer vielversprechenden Strategie zur Generierung von maßgeschneiderten Zufallsbibliotheken für Aminoacyl-tRNA Synthetasen hinsichtlich des Einbaus sterisch anspruchsvoller Aminosäure-Substrate sowie der Verbesserung der Amber-Suppression durch die Inaktivierung des Rf-1 zeigen das Potential der synthetischen Biologie auf dem Weg zur präparativen Anwendung.

6 Abkürzungsverzeichnis

AEX	Anionenaustauschchromatographie
Amp	Ampicillin
AP	Alkalische Phosphatase
aTc	Anhydrotetracyclin
ATP	Adenosin-5'-Triphosphat
AzAla	Azulenylalanin
Bfa	Borono-2,6-difluorphenyl-alanin
bp	Basenpaar
Bpa	<i>p</i> -Borono-L-phenylalanin
Cam	Chloramphenicol
dNTP	Desoxynucleosidtriphosphat
Hca	7-Hydroxycumarylalanin
HqAla	8-Hydroxychinolinyllalanin
IMAC	Metallchelat-Affinitätschromatographie
IPTG	Isopropyl- β -D-thiogalactopyranosid
kb	Kilobase
LB	Luria-Bertani
$M_{\text{kalk.}}$	Kalkulierte molare Masse
$M_{\text{obs.}}$	Gemessene molare Masse
OD _x	Optische Dichte bei Wellenlänge x
ori	Replikationsursprung
PAGE	Polyacrylamid-Gelelektrophorese
PAS	Polypeptid aus repetitiven Abfolgen der Aminosäuren Prolin, Alanin und Serin
PCR	Polymerase-Kettenreaktion
PhotoDopa	<i>o</i> -Nitrobenzyl-dopamin
Rf-1	Release Factor 1
Rf-2	Release Factor 2
rpm	Umdrehungen pro Minute
RT	Raumtemperatur
SAC	Streptavidin-Affinitätschromatographie
SDS	Natriumdodecylsulfat
SEC	Größenausschlusschromatographie

<i>Strep</i> -tag	Streptavidin-Affinitätsanhängsel
TB	<i>Terrific Broth</i>
v/v	Volumenprozent
w/v	Gewichtsprozent
wt	Wildtyp
ϵ	Molarer Extinktionskoeffizient

7 Literaturverzeichnis

- Ai, H.W. (2012) Biochemical analysis with the expanded genetic lexicon. *Anal. Bioanal. Chem.* 403, 2089-2102.
- Åkerström, B., Borregaard, N., Lögdberg, L., Flower, D. & Salier, J.-S. (2006) Lipocalins. *Landes Bioscience*. Georgetown, Texas.
- Ambrogelly, A., Palioura, S. & Söll, D. (2007) Natural expansion of the genetic code. *Nat. Chem. Biol.* 3, 29-35.
- Ambrosi, M., Cameron, N.R. & Davis, B.G. (2005) Lectins: tools for the molecular understanding of the glycode. *Org. Biomol. Chem.* 3, 1593-1608.
- Amiram, M., Haimovich, A.D., Fan, C., Wang, Y.S., Aerni, H.R., Ntai, I., Moonan, D.W., Ma, N.J., Rovner, A.J., Hong, S.H., Kelleher, N.L., Goodman, A.L., Jewett, M.C., Söll, D., Rinehart, J. & Isaacs, F.J. (2015) Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat. Biotechnol.* 33, 1272-1279.
- Anderson, J.C., Wu, N., Santoro, S.W., Lakshman, V., King, D.S. & Schultz, P.G. (2004) An expanded genetic code with a functional quadruplet codon. *Proc. Nat. Acad. Sci. U.S.A* 101, 7566-7571.
- Badran, A.H. & Liu, D.R. (2015) Development of potent in vivo mutagenesis plasmids with broad mutational spectra. *Nat. Commun.* 6, 8425.
- Baumann, T., Hauf, M., Richter, F., Albers, S., Möglich, A., Ignatova, Z. & Budisa, N. (2019a) Computational Aminoacyl-tRNA Synthetase Library Design for Photocaged Tyrosine. *Int J Mol Sci* 20.
- Baumann, T., Hauf, M., Schildhauer, F., Eberl, K.B., Durkin, P.M., Deniz, E., Löffler, J.G., Acevedo-Rocha, C.G., Jaric, J., Martins, B.M., Dobbek, H., Bredenbeck, J. & Budisa, N. (2019b) Site-Resolved Observation of Vibrational Energy Transfer Using a Genetically Encoded Ultrafast Heater. *Angew. Chem. Int. Ed. Engl.* 58, 2899-2903.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. (2000) The Protein Data Bank. *Nucleic. Acids Res.* 28, 235-242.
- Beyer, J.N., Hosseinzadeh, P., Gottfried-Lee, I., Van Fossen, E.M., Zhu, P., Bednar, R.M., Karplus, P.A., Mehl, R.A. & Cooley, R.B. (2020) Overcoming Near-Cognate Suppression in a Release Factor 1-Deficient Host with an Improved Nitro-Tyrosine tRNA Synthetase. *J. Mol. Biol.* 432, 4690-4704.
- Blanas, A., Sahasrabudhe, N.M., Rodríguez, E., van Kooyk, Y. & van Vliet, S.J. (2018) Fucosylated Antigens in Cancer: An Alliance toward Tumor Progression, Metastasis, and Resistance to Chemotherapy. *Front. Oncol.* 8, 39.
- Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. & Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453-1462.

- Böeseken, J. (1949) The use of boric acid for the determination of the configuration of carbohydrates. *Advances in carbohydrate chemistry*. Elsevier.
- Breustedt, D.A., Schönfeld, D.L. & Skerra, A. (2006) Comparative ligand-binding analysis of ten human lipocalins. *Biochim. Biophys. Acta* 1764, 161-173.
- Brick, P. & Blow, D.M. (1987) Crystal structure of a deletion mutant of a tyrosyl-tRNA synthetase complexed with tyrosine. *J. Mol. Biol.* 194, 287-297.
- Brun, M.P., Bischoff, L. & Garbay, C. (2004) A very short route to enantiomerically pure coumarin-bearing fluorescent amino acids. *Angew. Chem. Int. Ed. Engl.* 43, 3432-3436.
- Brustad, E., Bushey, M.L., Lee, J.W., Groff, D., Liu, W. & Schultz, P.G. (2008) A Genetically Encoded Boronate-Containing Amino Acid. *Angew. Chem., Int. Ed.* 47, 8220-8223.
- Budisa, N. (2006) *Engineering the genetic code: expanding the amino acid repertoire for the design of novel proteins*, John Wiley & Sons.
- Bullock, W.O. (1987) XL1-Blue: a high efficiency plasmid transforming recA Escherichia coli strain with beta-galactosidase selection. *Biotechniques* 5, 376-379.
- Burke, S.A., Lo, S.L. & Krzycki, J.A. (1998) Clustered genes encoding the methyltransferases of methanogenesis from monomethylamine. *J. Bacteriol* 180, 3432-3440.
- Carneiro, A., Matos, M.J., Uriarte, E. & Santana, L. (2021) Trending Topics on Coumarin and Its Derivatives in 2020. *Molecules* 26.
- Chatterjee, A., Guo, J., Lee, H.S. & Schultz, P.G. (2013) A genetically encoded fluorescent probe in mammalian cells. *J. Am. Chem. Soc.* 135, 12540-12543.
- Cheng, Z., Kuru, E., Sachdeva, A. & Vendrell, M. (2020) Fluorescent amino acids as versatile building blocks for chemical biology. *Nat. Rev. Chem.* 4, 275-290.
- Chien, A., Edgar, D.B. & Trela, J.M. (1976) Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J. Bacteriol.* 127, 1550-1557.
- Cobb, M. (2017) 60 years ago, Francis Crick changed the logic of biology. *PLoS Biol.* 15, e2003243.
- Cohen, S.N., Chang, A.C.Y. & Hsu, L. (1972) Nonchromosomal Antibiotic Resistance in Bacteria: Genetic Transformation of *Escherichia coli* by R-Factor DNA. *Proc. Nat. Acad. Sci. U.S.A.* 69, 2110-2114.
- Concordet, J.P. & Haeussler, M. (2018) CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic. Acids Res.* 46, W242-w245.
- Crick, F.H. (1958) On protein synthesis. *Symp. Soc. Exp. Biol.* 12, 138-163.
- Crick, F.H. (1968) The origin of the genetic code. *J. Mol. Biol.* 38, 367-379.
- Dalby, P.A. (2003) Optimising enzyme function by directed evolution. *Curr. Opin. Struct. Biol.* 13, 500-505.

- Datsenko, K.A. & Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Nat. Acad. Sci. U.S.A.* 97, 6640-6645.
- Daugherty, P.S., Iverson, B.L. & Georgiou, G. (2000) Flow cytometric screening of cell-based libraries. *J. Immunol. Methods* 243, 211-227.
- Davis, A.P. (2009) Synthetic lectins. *Org. Biomol. Chem.* 7, 3629-3638.
- Davis, L. & Chin, J.W. (2012) Designer proteins: applications of genetic code expansion in cell biology. *Nat. Rev. Mol. Cell. Biol.* 13, 168-182.
- DeLano, W.L. (2002) Pymol: An open-source molecular graphics tool. Palo Alto, CA, USA, Delano Scientific.
- Deuschle, F.-C., Ilyukhina, E. & Skerra, A. (2020) Anticalin®-Proteine als neue Wirkstoffklasse für therapeutische Anwendungen. *Pharm. Ind.* 82, 283-289.
- Duan, L., Zangiabadi, M. & Zhao, Y. (2020) Synthetic lectins for selective binding of glycoproteins in water. *Chem. Commun.* 56, 10199-10202.
- Edwardraja, S., Eichinger, A., Theobald, I., Sommer, C.A., Reichert, A.J. & Skerra, A. (2017) Rational Design of an Anticalin-Type Sugar-Binding Protein Using a Genetically Encoded Boronate Side Chain. *ACS Synth. Biol.* 6, 2241-2247.
- Eggertsson, G. & Söll, D. (1988) Transfer ribonucleic acid-mediated suppression of termination codons in *Escherichia coli*. *Microbiol Rev* 52, 354-374.
- Eichinger, A., Rauth, S., Hinz, D., Feuerbach, A. & Skerra, A. (2022) Structural basis of Alzheimer β -amyloid peptide recognition by engineered lipocalin proteins with aggregation-blocking activity. *Biol Chem* 403, 557-571.
- Falkowska, E., Le, K.M., Ramos, A., Doores, K.J., Lee, J.H., Blattner, C., Ramirez, A., Derking, R., van Gils, M.J., Liang, C.H., McBride, R., von Bredow, B., Shivatare, S.S., Wu, C.Y., Chan-Hui, P.Y., Liu, Y., Feizi, T., Zwick, M.B., Koff, W.C., Seaman, M.S., Swiderek, K., Moore, J.P., Evans, D., Paulson, J.C., Wong, C.H., Ward, A.B., Wilson, I.A., Sanders, R.W., Poignard, P. & Burton, D.R. (2014) Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. *Immunity.* 40, 657-668.
- Ferrige, A.G., Seddon, M.J., Jarvis, S., Skilling, J. & Aplin, R. (1991) Maximum entropy deconvolution in electrospray mass spectrometry. *Rapid Commun. Mass. Spectrom.* 5, 374-377.
- Fling, S.P. & Gregerson, D.S. (1986) Peptide and protein molecular weight determination by electrophoresis using a high-molarity tris buffer system without urea. *Anal. Biochem.* 155, 83-88.
- Forchhammer, K. & Böck, A. (1991) Biologie und Biochemie des Elements Selen. *Naturwissenschaften* 78, 497-504.

- Gabius, H.J., André, S., Jiménez-Barbero, J., Romero, A. & Solís, D. (2011) From lectin structure to functional glycomics: principles of the sugar code. *Trends Biochem. Sci.* **36**, 298-313.
- Galindo Casas, M., Stargardt, P., Mairhofer, J. & Wiltschi, B. (2020) Decoupling Protein Production from Cell Growth Enhances the Site-Specific Incorporation of Noncanonical Amino Acids in *E. coli*. *ACS Synth. Biol.* **9**, 3052-3066.
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D. & Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784-3788.
- Gebauer, M., Schiefner, A., Matschiner, G. & Skerra, A. (2013) Combinatorial Design of an Anticalin Directed against the Extra-Domain B for the Specific Targeting of Oncofetal Fibronectin. *J. Mol. Biol.* **425**, 780-802.
- Gebauer, M. & Skerra, A. (2012) Anticalins small engineered binding proteins based on the lipocalin scaffold. *Methods Enzymol.* **503**, 157-188.
- Gerdes, S.Y., Scholle, M.D., Campbell, J.W., Balázsi, G., Ravasz, E., Daugherty, M.D., Somera, A.L., Kyrpides, N.C., Anderson, I., Gelfand, M.S., Bhattacharya, A., Kapatral, V., D'Souza, M., Baev, M.V., Grechkin, Y., Mseeh, F., Fonstein, M.Y., Overbeek, R., Barabási, A.L., Oltvai, Z.N. & Osterman, A.L. (2003) Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J. Bacteriol.* **185**, 5673-5684.
- Gill, S.C. & von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.
- Hakomori, S. (1985) Aberrant glycosylation in cancer cell membranes as focused on glycolipids: overview and perspectives. *Cancer Res.* **45**, 2405-2414.
- Hao, B., Gong, W., Ferguson, T.K., James, C.M., Krzycki, J.A. & Chan, M.K. (2002) A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* **296**, 1462-1466.
- Hauf, M., Richter, F., Schneider, T., Faidt, T., Martins, B.M., Baumann, T., Durkin, P., Dobbek, H., Jacobs, K., Möglich, A. & Budisa, N. (2017) Photoactivatable Mussel-Based Underwater Adhesive Proteins by an Expanded Genetic Code. *ChemBiochem* **18**, 1819-1823.
- Hinegardner, R.T. & Engelberg, J. (1963) RATIONALE FOR A UNIVERSAL GENETIC CODE. *Science* **142**, 1083-1085.
- Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. & Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59.
- Hogrefe, H. (1997) Polymerase enhancing factor (PEF) extracts, PEF protein complexes, isolated PEF proteins, and methods for purifying and identifying same. Agilent Technologies Inc

- Hogrefe, H.H. & Cline, J.M. (2001) Compositions and methods for random nucleic acid mutagenesis Agilent Technologies Inc
- Hogrefe, H.H. & Hansen, C.J. (1997) Polymerase enhancing factor (PEF) extracts, PEF protein complexes, isolated PEF protein, and methods for purifying and identifying. Agilent Technologies Inc
- Inokuchi, H., Yamao, F., Sakano, H. & Ozeki, H. (1979) Identification of transfer RNA suppressors in *Escherichia coli*. I. Amber suppressor su+2, an anticodon mutant of tRNA^{2Gln}. *J. Mol. Biol.* 132, 649-662.
- Jakubowski, H. & Goldman, E. (1992) Editing of errors in selection of amino acids for protein synthesis. *Microbiol. Rev.* 56, 412-429.
- Jancarik, J. & Kim, S.-H. (1991) Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallogr.* 24, 409-411.
- Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J. & Yang, S. (2015) Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 81, 2506-2514.
- Jin, X., Park, O.J. & Hong, S.H. (2019) Incorporation of non-standard amino acids into proteins: challenges, recent achievements, and emerging applications. *Appl. Microbiol. Biotechnol.* 103, 2947-2958.
- Johnson, D.B., Wang, C., Xu, J., Schultz, M.D., Schmitz, R.J., Ecker, J.R. & Wang, L. (2012) Release factor one is nonessential in *Escherichia coli*. *ACS Chem. Biol.* 7, 1337-1344.
- Johnson, D.B., Xu, J., Shen, Z., Takimoto, J.K., Schultz, M.D., Schmitz, R.J., Xiang, Z., Ecker, J.R., Briggs, S.P. & Wang, L. (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol.* 7, 779-786.
- Kang, M., Lu, Y., Chen, S. & Tian, F. (2018) Harnessing the power of an expanded genetic code toward next-generation biopharmaceuticals. *Curr. Opin. Chem. Biol.* 46, 123-129.
- Kappler, K. & Hennet, T. (2020a) Emergence and significance of carbohydrate-specific antibodies. *Genes Immun.* 21, 224-239.
- Kappler, K. & Hennet, T. (2020b) Emergence and significance of carbohydrate-specific antibodies. *Genes Immun.* 21, 224-239.
- Kapust, R.B., Tözsér, J., Fox, J.D., Anderson, D.E., Cherry, S., Copeland, T.D. & Waugh, D.S. (2001) Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* 14, 993-1000.
- Kapust, R.B. & Waugh, D.S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 8, 1668-1674.
- Kers, J.A., DeFusco, A.W., Park, J.H., Xu, J., Pulse, M.E., Weiss, W.J. & Handfield, M. (2018) OG716: Designing a fit-for-purpose lantibiotic for the treatment of *Clostridium difficile* infections. *PLoS One* 13, e0197467.

- Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K. & Yokoyama, S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol.* 10, 425-432.
- Korostelev, A., Asahara, H., Lancaster, L., Laurberg, M., Hirschi, A., Zhu, J., Trakhanov, S., Scott, W.G. & Noller, H.F. (2008) Crystal structure of a translation termination complex formed with release factor RF2. *Proc. Nat. Acad. Sci. U.S.A.* 105, 19684-19689.
- Korostelev, A., Zhu, J., Asahara, H. & Noller, H.F. (2010) Recognition of the amber UAG stop codon by release factor RF1. *Embo J.* 29, 2577-2585.
- Krahn, N., Tharp, J.M., Crnković, A. & Söll, D. (2020) Engineering aminoacyl-tRNA synthetases for use in synthetic biology. *Enzymes* 48, 351-395.
- Kretz, K., Callen, W. & Hedden, V. (1994) Cycle sequencing. *PCR Methods Appl* 3, S107-112.
- Kubik, S. (2009) Synthetic lectins. *Angew. Chem. Int. Ed. Engl.* 48, 1722-1725.
- Kuhn, S.M., Rubini, M., Fuhrmann, M., Theobald, I. & Skerra, A. (2010) Engineering of an Orthogonal Aminoacyl-tRNA Synthetase for Efficient Incorporation of the Non-natural Amino Acid O-Methyl-L-tyrosine using Fluorescence-based Bacterial Cell Sorting. *J. Mol. Biol.* 404, 70-87.
- Kuhn, S.M., Rubini, M., Müller, M.A. & Skerra, A. (2011) Biosynthesis of a fluorescent protein with extreme pseudo-Stokes shift by introducing a genetically encoded non-natural amino acid outside the fluorophore. *J. Am. Chem. Soc.* 133, 3708-3711.
- Kuznetsov, G., Goodman, D.B., Filsinger, G.T., Landon, M., Rohland, N., Aach, J., Lajoie, M.J. & Church, G.M. (2017) Optimizing complex phenotypes through model-guided multiplex genome engineering. *Genome Biol.* 18, 100.
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Laine, R.A. (1994) A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* 4, 759-767.
- Lajoie, M.J., Rovner, A.J., Goodman, D.B., Aerni, H.R., Haimovich, A.D., Kuznetsov, G., Mercer, J.A., Wang, H.H., Carr, P.A., Mosberg, J.A., Rohland, N., Schultz, P.G., Jacobson, J.M., Rinehart, J., Church, G.M. & Isaacs, F.J. (2013) Genomically recoded organisms expand biological functions. *Science* 342, 357-360.
- Lakowicz, J.R. (2013) *Principles of fluorescence spectroscopy*, Springer science & business media.
- Lang, K., Davis, L. & Chin, J.W. (2015) Genetic encoding of unnatural amino acids for labeling proteins. *Methods Mol. Biol.* 1266, 217-228.
- Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S. & Noller, H.F. (2008) Structural basis for translation termination on the 70S ribosome. *Nature* 454, 852-857.

- Lei, J., Hansen, G., Nitsche, C., Klein, C.D., Zhang, L. & Hilgenfeld, R. (2016) Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. *Science* 353, 503-505.
- Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M.A. & Böck, A. (1988) Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature* 331, 723-725.
- Lemmerer, M., Mairhofer, J., Lepak, A., Longus, K., Hahn, R. & Nidetzky, B. (2019) Decoupling of recombinant protein production from Escherichia coli cell growth enhances functional expression of plant Leloir glycosyltransferases. *Biotechnol. Bioeng.* 116, 1259-1268.
- Leney, A.C. & Heck, A.J. (2017) Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* 28, 5-13.
- Liu, D.R. & Schultz, P.G. (1999) Progress toward the evolution of an organism with an expanded genetic code. *Proc. Nat. Acad. Sci. U.S.A.* 96, 4780-4785.
- Liu, X., Li, J., Hu, C., Zhou, Q., Zhang, W., Hu, M., Zhou, J. & Wang, J. (2013) Significant expansion of the fluorescent protein chromophore through the genetic incorporation of a metal-chelating unnatural amino acid. *Angew. Chem. Int. Ed. Engl.* 52, 4805-4809.
- Melnikov, S.V. & Söll, D. (2019) Aminoacyl-tRNA Synthetases and tRNAs for an Expanded Genetic Code: What Makes them Orthogonal? *Int. J. Mol. Sci.* 20.
- Morrison, K.L. & Weiss, G.A. (2001) Combinatorial alanine-scanning. *Curr. Opin. Chem. Biol.* 5, 302-307.
- Movva, N.R., Nakamura, K. & Inouye, M. (1980) Amino acid sequence of the signal peptide of ompA protein, a major outer membrane protein of Escherichia coli. *J. Biol. Chem.* 255, 27-29.
- Mukai, T., Hayashi, A., Iraha, F., Sato, A., Ohtake, K., Yokoyama, S. & Sakamoto, K. (2010) Codon reassignment in the Escherichia coli genetic code. *Nucleic Acids Res.* 38, 8188-8195.
- Murin, C.D., Wilson, I.A. & Ward, A.B. (2019) Antibody responses to viral infections: a structural perspective across three different enveloped viruses. *Nat. Microbiol.* 4, 734-747.
- Nakamura, K. & Inouye, M. (1979) DNA sequence of the gene for the outer membrane lipoprotein of E. coli: an extremely AT-rich promoter. *Cell* 18, 1109-1117.
- Nakamura, Y., Ito, K., Matsumura, K., Kawazu, Y. & Ebihara, K. (1995) Regulation of translation termination: conserved structural motifs in bacterial and eukaryotic polypeptide release factors. *Biochem. Cell Biol.* 73, 1113-1122.
- Nene, V. & Glass, R.E. (1981) Direct measurement of absolute suppressor efficiency. *Biosci Rep* 1, 571-574.

- Neumann, H. (2012) Rewiring translation - Genetic code expansion and its applications. *FEBS Lett.* 586, 2057-2064.
- Noble, P., Spendlove, I., Harding, S., Parsons, T. & Durrant, L.G. (2013) Therapeutic targeting of Lewis(y) and Lewis(b) with a novel monoclonal antibody 692/29. *PLoS One* 8, e54892.
- Ongey, E.L., Yassi, H., Pflugmacher, S. & Neubauer, P. (2017) Pharmacological and pharmacokinetic properties of lanthipeptides undergoing clinical studies. *Biotechnol. Lett.* 39, 473-482.
- Pang, Y.L., Poruri, K. & Martinis, S.A. (2014) tRNA synthetase: tRNA aminoacylation and beyond. *Wiley Interdiscip. Rev. RNA* 5, 461-480.
- Parren, P.W.H.I. & Lugovskoy, A.A. (2013) Therapeutic antibody engineering. *mAbs* 5, 175-177.
- Pastrnak, M., Magliery, T.J. & Schultz, P.G. (2000) A New Orthogonal Suppressor tRNA/Aminoacyl-tRNA Synthetase Pair for Evolving an Organism with an Expanded Genetic Code. *Helv. Chim. Acta* 83, 2277-2286.
- Peters, J.A. (2014) Interactions between boric acid derivatives and saccharides in aqueous media: Structures and stabilities of resulting esters. *Coord. Chem. Rev.* 268, 1-22.
- Pinto, D., Park, Y.J., Beltramello, M., Walls, A.C., Tortorici, M.A., Bianchi, S., Jaconi, S., Culap, K., Zatta, F., De Marco, A., Peter, A., Guarino, B., Spreafico, R., Cameroni, E., Case, J.B., Chen, R.E., Havenar-Daughton, C., Snell, G., Telenti, A., Virgin, H.W., Lanzavecchia, A., Diamond, M.S., Fink, K., Veesler, D. & Corti, D. (2020) Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* 583, 290-295.
- Pizer, R. & Babcock, L. (1977) Mechanism of the complexation of boron acids with catechol and substituted catechols. *Inorg. Chem.* 16, 1677-1681.
- Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 258, 598-599.
- Prescher, J.A. & Bertozzi, C.R. (2005) Chemistry in living systems. *Nat. Chem. Biol.* 1, 13-21.
- Reichert, A.J., Poxleitner, G., Dauner, M. & Skerra, A. (2015) Optimisation of a system for the co-translational incorporation of a keto amino acid and its application to a tumour-specific Anticalin. *Protein Eng. Des. Sel.* 28, 553-565.
- Reily, C., Stewart, T.J., Renfrow, M.B. & Novak, J. (2019) Glycosylation in health and disease. *Nat. Rev. Nephrol.* 15, 346-366.
- Richter, A., Eggenstein, E. & Skerra, A. (2014) Anticalins: exploiting a non-Ig scaffold with hypervariable loops for the engineering of binding proteins. *FEBS Lett.* 588, 213-218.
- Riggs, A.D., Itakura, K., Crea, R., Hirose, T., Kraszewski, A., Goeddel, D., Kleid, D., Yansura, D.G., Bolivar, F. & Heyneker, H.L. (1980) Synthesis, cloning, and expression of hormone genes in Escherichia coli. *Recent Prog Horm Res* 36, 261-276.

- Rothe, C. & Skerra, A. (2018) Anticalin(®) Proteins as Therapeutic Agents in Human Diseases. *BioDrugs* 32, 233-243.
- Rouvinski, A., Guardado-Calvo, P., Barba-Spaeth, G., Duquerroy, S., Vaney, M.C., Kikuti, C.M., Navarro Sanchez, M.E., Dejnirattisai, W., Wongwiwat, W., Haouz, A., Girard-Blanc, C., Petres, S., Shepard, W.E., Desprès, P., Arenzana-Seisdedos, F., Dussart, P., Mongkolsapaya, J., Sreaton, G.R. & Rey, F.A. (2015) Recognition determinants of broadly neutralizing human antibodies against dengue viruses. *Nature* 520, 109-113.
- Rydén, S.M. & Isaksson, L.A. (1984) A temperature-sensitive mutant of Escherichia coli that shows enhanced misreading of UAG/A and increased efficiency for some tRNA nonsense suppressors. *Mol. Gen. Genet.* 193, 38-45.
- Sadr, V., Saffar, B. & Emamzadeh, R. (2017) Functional expression and purification of recombinant Hepcidin25 production in Escherichia coli using SUMO fusion technology. *Gene* 610, 112-117.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. & Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239, 487-491.
- Saleh, M.N., Sugarman, S., Murray, J., Ostroff, J.B., Healey, D., Jones, D., Daniel, C.R., LeBherz, D., Brewer, H., Onetto, N. & LoBuglio, A.F. (2000) Phase I trial of the anti-Lewis Y drug immunoconjugate BR96-doxorubicin in patients with lewis Y-expressing epithelial tumors. *J. Clin. Oncol.* 18, 2282-2292.
- Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.
- Sanger, F., Nicklen, S. & Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci. U.S.A.* 74, 5463-5467.
- Schiefner, A. & Skerra, A. (2015) The menagerie of human lipocalins: a natural protein scaffold for molecular recognition of physiological compounds. *Acc. Chem. Res.* 48, 976-985.
- Schira, K. (2020) Fakultative Inaktivierung des Release Factor 1 in E. coli für den verbesserten Einbau von künstlichen Aminosäuren in rekombinante Proteine durch Amber-Suppression. *Lehrstuhl für Biologische Chemie*. TU München.
- Schmidt, T.G. & Skerra, A. (1994) One-step affinity purification of bacterially produced proteins by means of the "Strep tag" and immobilized recombinant core streptavidin. *J. Chromatogr. A.* 676, 337-345.
- Schmidt, T.G.M. & Skerra, A. (2007) The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protocols* 2, 1528-1535.
- Scolnick, E., Tompkins, R., Caskey, T. & Nirenberg, M. (1968) Release factors differing in specificity for terminator codons. *Proc. Nat. Acad. Sci. U.S.A.* 61, 768-774.
- Scott, A.M., Tebbutt, N., Lee, F.T., Cavicchiolo, T., Liu, Z., Gill, S., Poon, A.M., Hopkins, W., Smyth, F.E., Murone, C., MacGregor, D., Papenfuss, A.T., Chappell, B., Saunder, T.H., Brechbiel, M.W., Davis, I.D., Murphy, R., Chong, G., Hoffman, E.W. & Old, L.J.

- (2007) A phase I biodistribution and pharmacokinetic trial of humanized monoclonal antibody Hu3s193 in patients with advanced epithelial cancers that express the Lewis-Y antigen. *Clin. Cancer Res.* 13, 3286-3292.
- Seifert, M.H., Ksiazek, D., Azim, M.K., Smialowski, P., Budisa, N. & Holak, T.A. (2002) Slow exchange in the chromophore of a green fluorescent protein variant. *J Am Chem Soc* 124, 7932-7942.
- Sharan, S.K., Thomason, L.C., Kuznetsov, S.G. & Court, D.L. (2009) Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.* 4, 206-223.
- Shaw, J.J. & Green, R. (2007) Two distinct components of release factor function uncovered by nucleophile partitioning analysis. *Mol. Cell.* 28, 458-467.
- Silverstone, A.E., Arditti, R.R. & Magasanik, B. (1970) Catabolite-insensitive revertants of lac promoter mutants. *Proc. Nat. Acad. Sci. U.S.A.* 66, 773-779.
- Skerra, A. (1994) Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene* 151, 131-135.
- Skerra, A. (2000) Lipocalins as a scaffold. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* 1482, 337-350.
- Skerra, A., Pfitzinger, I. & Plückthun, A. (1991) The functional expression of antibody Fv fragments in *Escherichia coli*: improved vectors and a generally applicable purification technique. *Nat. Biotechnol.* 9, 273-278.
- Skerra, A. & Plückthun, A. (1988) Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. *Science* 240, 1038-1041.
- Skerra, A., Reichert, A.J. & Dauner, M. (2018) Light-switchable polypeptide and uses thereof. TU München.
- Skerra, A. & Schmidt, T.G. (2000) Use of the Strep-Tag and streptavidin for detection and purification of recombinant proteins. *Methods Enzymol.* 326, 271-304.
- Sommer, C.A., Eichinger, A. & Skerra, A. (2020) A Tetrahedral Boronic Acid Diester Formed by an Unnatural Amino Acid in the Ligand Pocket of an Engineered Lipocalin. *ChemBioChem* 21, 469-472.
- Springsteen, G. & Wang, B. (2002) A detailed examination of boronic acid–diol complexation. *Tetrahedron* 58, 5291-5300.
- Srinivasan, G., James, C.M. & Krzycki, J.A. (2002) Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* 296, 1459-1462.
- Stemmer, W.P. (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Nat. Acad. Sci. U.S.A.* 91, 10747-10751.
- Studier, F.W. & Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113-130.

- Summerer, D., Chen, S., Wu, N., Deiters, A., Chin, J.W. & Schultz, P.G. (2006) A genetically encoded fluorescent amino acid. *Proc. Nat. Acad. Sci. U.S.A.* 103, 9785-9789.
- Tian, F., Lu, Y., Manibusan, A., Sellers, A., Tran, H., Sun, Y., Phuong, T., Barnett, R., Hehli, B., Song, F., DeGuzman, M.J., Ensari, S., Pinkstaff, J.K., Sullivan, L.M., Biroc, S.L., Cho, H., Schultz, P.G., DiJoseph, J., Dougher, M., Ma, D., Dushin, R., Leal, M., Tchistiakova, L., Feyfant, E., Gerber, H.P. & Sapra, P. (2014) A general approach to site-specific antibody drug conjugates. *Proc. Nat. Acad. Sci. U.S.A.* 111, 1766-1771.
- Toone, E.J. (1994) Structure and energetics of protein-carbohydrate complexes. *Curr. Opin. Struct. Biol.* 4, 719-728.
- Toseland, C.P. (2013) Fluorescent labeling and modification of proteins. *J. Chem. Biol.* 6, 85-95.
- Tseng, H.W., Baumann, T., Sun, H., Wang, Y.S., Ignatova, Z. & Budisa, N. (2020) Expanding the Scope of Orthogonal Translation with Pyrrolysyl-tRNA Synthetases Dedicated to Aromatic Amino Acids. *Molecules* 25.
- Tung, W.L. & Chow, K.C. (1995) A modified medium for efficient electrotransformation of *E. coli*. *Trends Genet.* 11, 128-129.
- Ulmer, K.M. (1983) Protein engineering. *Science* 219, 666-671.
- Umehara, T., Kim, J., Lee, S., Guo, L.T., Söll, D. & Park, H.S. (2012) N-acetyl lysyl-tRNA synthetases evolved by a CcdB-based selection possess N-acetyl lysine specificity in vitro and in vivo. *FEBS Lett.* 586, 729-733.
- Uno, M., Ito, K. & Nakamura, Y. (1996) Functional specificity of amino acid at position 246 in the tRNA mimicry domain of bacterial release factor 2. *Biochimie* 78, 935-943.
- v. Vargha, L. (1933) Partielle Acetonierung der Zucker und Zuckeralkohole, I. Mitteil: 1.2-Aceton-d-glucose-3.5-monoborsäureester(1.4). *Berichte der deutschen chemischen Gesellschaft (A and B Series)* 66, 704-707.
- van den Berg, R., Peters, J.A. & van Bakkum, H. (1994) The structure and (local) stability constants of borate esters of mono- and di-saccharides as studied by ¹¹B and ¹³C NMR spectroscopy. *Carbohydr. Res.* 253, 1-12.
- Van den Brulle, J., Fischer, M., Langmann, T., Horn, G., Waldmann, T., Arnold, S., Fuhrmann, M., Schatz, O., O'Connell, T., O'Connell, D., Auckenthaler, A. & Schwer, H. (2008) A novel solid phase technology for high-throughput gene synthesis. *BioTechniques* 45, 340-343.
- Varki, A. (2016) Biological roles of glycans. *Glycobiology* 27, 3-49.
- Vega, A., Zarate, M., Tlahuext, H. & Höpfl, H. (2010) Crystal structure of aqua-2-(3-aminophenyl)-1, 3, 2-dioxaborole dihydrate, C₁₂H₁₂BNO₃. 2H₂O. *Z. Kristallogr. - New Cryst. Struct.* 225, 676.
- Volkov, A.A. & Arnold, F.H. (2000) Methods for in vitro DNA recombination and random chimeragenesis. *Methods Enzymol.* 328, 447-456.

- Wals, K. & Ovaa, H. (2014) Unnatural amino acid incorporation in E. coli: current and future applications in the design of therapeutic proteins. *Front. Chem.* 2, 15.
- Wan, W., Tharp, J.M. & Liu, W.R. (2014) Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim. Biophys. Acta* 1844, 1059-1070.
- Wang, D., Tang, J., Tang, J. & Wang, L.X. (2015) Targeting N-glycan cryptic sugar moieties for broad-spectrum virus neutralization: progress in identifying conserved molecular targets in viruses of distinct phylogenetic origins. *Molecules* 20, 4610-4622.
- Wang, J., Xie, J. & Schultz, P.G. (2006) A genetically encoded fluorescent amino acid. *J. Am. Chem. Soc.* 128, 8738-8739.
- Wang, K., Sachdeva, A., Cox, D.J., Wilf, N.M., Lang, K., Wallace, S., Mehl, R.A. & Chin, J.W. (2014) Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* 6, 393-403.
- Wang, L., Brock, A., Herberich, B. & Schultz, P.G. (2001) Expanding the genetic code of Escherichia coli. *Science* 292, 498-500.
- Wang, L., Magliery, T.J., Liu, D.R. & Schultz, P.G. (2000) A New Functional Suppressor tRNA/Aminoacyl-tRNA Synthetase Pair for the in Vivo Incorporation of Unnatural Amino Acids into Proteins. *J. Am. Chem. Soc.* 122, 5010-5011.
- Wang, L. & Schultz, P.G. (2001) A general approach for the generation of orthogonal tRNAs. *Chem. Biol.* 8, 883-890.
- Weis, W.I. & Drickamer, K. (1996) Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* 65, 441-473.
- Wiedenheft, B., Sternberg, S.H. & Doudna, J.A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331-338.
- Woese, C.R. (1964) UNIVERSALITY IN THE GENETIC CODE. *Science* 144, 1030-1031.
- Woese, C.R. (1965) Order in the genetic code. *Proc. Nat. Acad. Sci. U.S.A.* 54, 71-75.
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103-119.
- Yin, G., Stephenson, H.T., Yang, J., Li, X., Armstrong, S.M., Heibeck, T.H., Tran, C., Masikat, M.R., Zhou, S., Stafford, R.L., Yam, A.Y., Lee, J., Steiner, A.R., Gill, A., Penta, K., Pollitt, S., Baliga, R., Murray, C.J., Thanos, C.D., McEvoy, L.M., Sato, A.K. & Hallam, T.J. (2017) RF1 attenuation enables efficient non-natural amino acid incorporation for production of homogeneous antibody drug conjugates. *Sci. Rep.* 7, 3026.
- Young, T.S. & Schultz, P.G. (2010) Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J. Biol. Chem.* 285, 11039-11044.

- Youngman, E.M., McDonald, M.E. & Green, R. (2008) Peptide release on the ribosome: mechanism and implications for translational control. *Annu. Rev. Microbiol.* 62, 353-373.
- Zeng, Y., Wang, W. & Liu, W.R. (2014) Towards reassigning the rare AGG codon in *Escherichia coli*. *Chembiochem* 15, 1750-1754.
- Zhang, M.S., Brunner, S.F., Huguenin-Dezot, N., Liang, A.D., Schmied, W.H., Rogerson, D.T. & Chin, J.W. (2017) Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nat Methods* 14, 729-736.
- Zheng, L., Baumann, U. & Reymond, J.L. (2004) An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* 32, e115.
- Zinoni, F., Birkmann, A., Leinfelder, W. & Böck, A. (1987) Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon. *Proc. Natl. Acad. Sci. USA* 84, 3156-3160.